# Student Answer Key

## Odd Questions Only

Ron C. Mittelhammer

# Mathematical Statistics for Economics and Business

*Second Edition*


For information contact Ron C. Mittelhammer, WSU, mittelha@wsu.edu
For corrections, please contact Danielle Engelhardt dkbishop@wsu.edu

Springer

**Chapter 1 – Student Answer Key – Odd Numbers Only**
**Elements of Probability Theory**

1. a) {0, 1, 2, ..., 100}

   b) {0, 1, 2, ..., $N$}, $N$ = number of ads circulated

   c) $\{(x, y): x \geq y,\ x \text{ and } y \in [-459.67,\ 212]\}$ (while perhaps conservative, the temperature range has been designated to be between absolute zero and the boiling point).

   d) $\{x : 0 \leq x \leq c\}$, where $c$ is delivery capacity of the jobber in a given week.

   e) {$w$: $w$ = \$$x.yz$, $x$ is a nonnegative integer, $y$ and $z$ ϵ{0, 1, 2,..., 9}}.

   f) {$x$: $x$ is a nonnegative integer}

3. a) $A$ = {Karen, Wendy, Brenda}
   $P(A) = 3/8$.

   b) $A$ = {Tom, Richard}
   $P(A) = 1/4$.

   c) $A$ = {Karen, Frank, Scott}
   $P(A) = 3/8$.

   d) $A$ = {Frank, Scott}
   $P(A) = 1/4$.

   e) $A = \{\emptyset\}$
   $P(A) = 0$.

   f) $A$ = {Tom, Karen, Frank, Eric, Wendy, Brenda, Scott, Richard}
   $P(A) = 1$.

5. Let $S = \{(x_1, x_2, x_3): x_1, x_2, x_3$ are three jokes chosen randomly from an inventory of 12 jokes$\}$.

   $S$ is a finite sample space, has equally likely outcomes, and

$$N(S) = \frac{12!}{3!9!} = 220,$$

represents the different combination of 12 jokes taken 3 at a time.

a) We need to count the number of ways in which 2 or all three jokes can be different from the previous month's three jokes, say $\left(x_1^0, x_2^0, x_3^0\right)$.

Let $A$ = 2 jokes are different. Then $N(A) = 3 \cdot \binom{9}{2} = 108$.

Let $B$ = all three are different. Then $N(B) = \binom{9}{3} = 84$.

So, $P\left(A \cup B\right) = \dfrac{N(A \cup B)}{220} = \dfrac{192}{220} = .8727.$

b) $P\left(B\right) = \dfrac{N(B)}{220} = \dfrac{84}{220} = .3818.$

7. **NOTE**: Each staff member has violated the axioms and theorems of probability, and there is more than one way to characterize the violation. In each case we identify one such characterization.

Tom: $\begin{array}{ccccc} P\left(A_1 \cap A_2\right) & \geq & P\left(A_1\right) & \text{or} & P\left(A_2\right) \\ .9 & \geq & .5 & \text{or} & .3 \end{array}.$

Dick: $\begin{array}{ccc} P\left(A_1 \cap A_2 \cap A_3\right) & \geq & 1 \\ 1.5 & \geq & 1 \end{array}.$

Harry: $\begin{array}{cc} P\left(A_3\right) & < 0 \\ -.4 & < 0 \end{array}.$

Sally: $P\left(A_3 \mid A_1\right) = \dfrac{P(A_3 \cap A_1)}{P(A_1)} = \dfrac{.3}{.2} = 1.5 > 1.$

9. .05(1000) = 50 chips will be inspected and there are 5 detectives in total.

Let $A_i$ = {event of picking a nondefective chip on $i^{\text{th}}$ draw}

$$P(\text{accepting shipment}) = P(A_1)\,P(A_2 \mid A_1) \prod_{i=3}^{50} P\left(A_i \mid \bigcap_{j=1}^{i-1} A_j\right),$$

$$= \frac{995}{1000} \bullet \frac{994}{999} \bullet \ldots \bullet \frac{946}{951} = \left(\frac{995!}{945!}\right) \div \left(\frac{1000!}{950!}\right) = \left(\frac{950!}{945!}\right) \div \left(\frac{1000!}{995!}\right),$$

$$= .7734.$$

11. To show Axioms 1-3 are satisfied: $P(A \mid B) \geq 0 \;\forall A \in \Upsilon_B$ since $P(A \cap B) \geq 0$, $P(B) > 0$, and thus $P(A \cup B)/P(B) \geq 0$.

$P(B \mid B) = P(B \cap B)/P(B) = P(B)/P(B) = 1.$

Let $A_i, i \in I$ be disjoint subsets of $B$. Then

$$P\left(\bigcup_{i\in I} A_i \mid B\right) = \frac{P\left(\left(\bigcup_{i\in I} A_i\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i\in I} (A_i \cap B)\right)}{P(B)}$$

$$= \frac{\sum_{i\in I} P(A_i \cap B)}{P(B)} = \sum_{i\in I} P(A_i \mid B).$$

since $(A_i \cap B), i \in I$, are disjoint events.

13. Let $D$ = {secretary dissatisfied}
$A_1$ = {dislikes supervisor}
$A_2$ = {not paid enough}
$A_3$ = {dislikes type of work}
$A_4$ = {conflict with other employees}

Note: $A_i \subset D$, $A_i \cap A_j = \varnothing$ for $i \neq j$, and $\cup_{i=1}^{4} A_i = D$, so the $A_i$'s form a partition of $D$.

Let $Q \subset D$ be the event that a dissatisfied secretary quits her job.

We are given the following:

| | |
|---|---|
| $P(D) = .20$ | $P(Q \mid A_1) = .20$ |
| $P(A_1 \mid D) = .55$ | $P(Q \mid A_2) = .30$ |
| $P(A_2 \mid D) = .30$ | $P(Q \mid A_3) = .90$ |
| $P(A_3 \mid D) = .10$ | $P(Q \mid A_4) = .05$ |
| $P(A_4 \mid D) = .05$ | |

a) Need to determine the $P(A_i \mid Q)$ values; therefore, using Bayes rule, we have:

$$P(A_i \mid Q) = \frac{P(Q \mid A_i)P(A_i)}{\sum_{i=1}^{4} P(Q \mid A_i)P(A_i)},$$

Now, $P(A_i \mid D) = \dfrac{P(A_i \cap D)}{P(D)} = \dfrac{P(A_i)}{P(D)} \Rightarrow P(A_i) = P(A_i \mid D)P(D)$, allowing us to rewrite Bayes rule as:

$$P(A_i \mid Q) = \frac{P(Q \mid A_i)P(A_i \mid D)P(D)}{\sum_{i=1}^{4} P(Q \mid A_i)P(A_i \mid D)P(D)},$$

where $\sum_{i=1}^{4} P(Q \mid A_i)P(A_i \mid D) = (.20)(.55) + (.30)(.30) + (.90)(.10) + (.05)(.05) = .2925.$

Then

$$P(A_1 \mid Q) = \frac{P(Q \mid A_1)P(A_1 \mid D)}{.2925} = \frac{(.20)(.55)}{.2925} = .3761$$

$$P(A_2 \mid Q) = \frac{(.30)(.30)}{.2925} = .3077$$ $\therefore$ The most probable reason is that she disliked her supervisor.

$$P(A_3 \mid Q) = \frac{(.90)(.10)}{.2925} = .3077$$

$$P(A_4 \mid Q) = \frac{(.05)(.05)}{.2925} = .0085$$

b) $A_2 \subset D \Rightarrow P(A_2) = P(A_2 \mid D)P(D) = (.30)(.20) = .06.$

c)

$$\begin{aligned}
P(Q \mid D) &= P(Q \cap D)/P(D) \\
&= P\left(Q \cap \left(\bigcup_{i=1}^{4} A_i\right)\right) / P(D), \\
&= \sum_{i=1}^{4} \frac{P(Q \cap A_i)}{P(D)} \quad \text{(since } \cap \text{ distributive and } Q \cap A_i, i = 1,\ldots,4 \text{ are disjoint)}, \\
&= \sum_{i=1}^{4} \frac{P(Q \cap A_i)}{P(A_i)} \frac{P(A_i)}{P(D)}, \\
&= \sum_{i=1}^{4} P(Q \mid A_i)P(A_i \mid D) \quad \left(\text{since } P(A_i) = P(A_i \mid D)P(D)\right), \\
&= .2925.
\end{aligned}$$

15. a) Yes. Using DeMorgan's Laws, note that $P(\bar{A}_1 \cap \bar{A}_2) = .18 = P(\bar{A}_1)P(\bar{A}_2)$,

$P(\bar{A}_1 \cap \bar{A}_3) = .2475 = P(\bar{A}_1)P(\bar{A}_3)$, and $P(\bar{A}_2 \cap \bar{A}_3) = .22 = P(\bar{A}_2)P(\bar{A}_3)$, so that Theorem 1.12 can then be applied.

b) No. Note that $P(A_2 \cap A_3 \mid A_1) = .20$, but $P(A_2 \cap A_3) = P(A_2) + P(A_3) - P(A_2 \cup A_3) = .60 + .45 - .78 = .27$, so conditional and unconditional probabilities are different.

c)

$$P(A_1 \cap A_2 \mid A_3) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_3)} = \frac{P(A_2 \cap A_3 \mid A_1)P(A_1)}{P(A_3)}$$

$$= \frac{(.20)(.55)}{.45} \qquad = .244$$

d) Yes, different since $P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) = .33$. Theorem 1.5 can be extended by letting $A = A_1$, $B = (A_2 \cup A_3)$, $\Rightarrow P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$ $- P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$,

or

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = {} & P(A_1) + P(A_2) + P(A_3) - [P(A_1) + P(A_2) - P(A_1 \cup A_2)] \\ & - [P(A_1) + P(A_3) - P(A_1 \cup A_3)] - [P(A_2) + P(A_3) - P(A_2 \cup A_3)] \\ & + P(A_1 \cap A_2 \cap A_3) \\ = {} & P(A_1 \cap A_2 \cap A_3) + P(A_1 \cup A_2) + P(A_2 \cup A_3) + P(A_1 \cup A_3) \\ & - P(A_1) - P(A_2) - P(A_3) \\ = {} & .11 + .82 + .78 + .7525 - .55 - .60 - .45 \\ = {} & .8625. \end{aligned}$$

17. $P(B|I) = .04$, $P(B|II) = .01$.

$P(I) = .85$, $P(II) = .15$.

a)

$$P(II \mid B) = \frac{P(B \mid II)P(II)}{P(B \mid II)\,P(II) + P(B \mid I)P(I)} \quad \text{(Bayes Rule)}$$

$$= \frac{(.01)(.15)}{(.01)(.15) + (.04)(.85)} = .04225.$$

b) $P(\bar{A} \mid B) = 1 - P(A \mid B) = 1 - .985 = .015.$

c)

$$P(B\mid A)=\frac{P(A\mid B)P(B)}{P(A\mid B)P(B)+P(A\mid \bar{B})P(\bar{B})} \quad \text{(Bayes Rule)}$$

$$=\frac{(.985)(.04)}{(.985)(.04)+(.015)(.96)}=.73234$$

The testing device is not especially accurate.

19. a)

$P(A\cap B)=.20,$

$P(A)P(B)=(.40)(.40)=.16,$

$P(A\cap B)\neq P(A)P(B)\Rightarrow A$ and $B$ are not pairwise independent.

$P(A\cap C)=.15$

$P(A)P(C)=(.40)(.375)=.15$

$P(A\cap C)=P(A)\ \ P(C)\Rightarrow A$ and $C$ are pairwise independent.

$P(B\cap C)=.15$

$P(B)P(C)=(.40)(.375)=.15$

$P(B\cap C)=P(B)\ \ P(C)\Rightarrow B$ and $C$ are pairwise independent.

b) No, pairwise independence is a necessary (not sufficient) condition for joint independence. Since $A$ and $B$ are not pairwise independent, we know immediately that $A$, $B$, and $C$ are not jointly independent.

c) $P(A\cap B) = .20$

$$P\big((A\cap B)\mid C\big)=\frac{P((A\cap B)\cap C)}{P(C)}=\frac{.10}{.375}=.267$$

d) Given $D\cap(A\cup B\cup C)=\varnothing\Rightarrow D\cap A=\varnothing$, then $P(D\cap A)=0$. But $P(D)P(A)=(.05)(.40)=.02\neq 0\Rightarrow D$ and $A$ are not independent. Given $P\big(D\cap(A\cup B\cup C)\big)=0$, then $P(D)P(A\cup B\cup C)=(.05)(.775)=.03875\neq 0\Rightarrow D$ and $(A\cup B\cup C)$ are not independent.

e) $$P\big(C\mid(A\cap B)\big)=\frac{P(C\cap(A\cap B))}{P(A\cap B)}=\frac{.10}{.20}=.50$$

21. a) The numerator of this function will always be $\geq 0$ for any $A \subset S$ (it will equal 0 when $A = \emptyset$). The denominator is always positive. This implies $P(A) \geq 0$, for any event $A \subset S$, thus Axiom 1.1 holds.

$P(S) = (1 + 2 + 3 + 4 + 5 + 6 + 7 + 8)/36 = 1$, thus Axiom 1.2 holds.

Let $\{A_i: i \in I\}$ be a collection of disjoint events contained in $S$. Then

$$P\left[\bigcup_{i\in I} A_i\right] = \sum_{x\in\left[\bigcup_{i\in I} A_i\right]} \left(x/36\right) = \sum_{i\in I}\sum_{x\in A_i} (x/36) \text{ (since } A_i\text{'s disjoint)} = \sum_{i\in I} P(A_i),$$

$\Rightarrow$ Axiom 1.3 holds .

All three axioms hold implying this is a valid probability set function.

b) The integrand $e^{-x} > 0 \forall x$, so that the integral of $e^{-x}$ over any event in the event space must be $\geq 0$, which implies Axiom 1.1 holds.

$$S = [0,\infty), \text{ so } P(S) = \int_0^\infty e^{-x}dx = \lim_{n\to\infty}(-e^{-n} + e^0) = 1 + \lim_{n\to\infty}\left(-e^{-n}\right) = 1$$

$\Rightarrow$ Axiom 1.2 holds .

$$P\left[\bigcup_{i\in I} A_i\right] = \int_{x\in\left[\bigcup_{i\in I} A_i\right]} e^{-x}\, dx = \sum_{i\in I}\int_{x\in A_i} e^{-x}dx \text{ (valid given the } A_i\text{'s are disjoint)} = \sum_{i\in I} P\left(A_i\right)$$

$\Rightarrow$ Axiom 1.3 holds .

All three axioms hold implying this is a valid probability set function.

c) This is not a probability set function. To see why, consider Axiom 1.2, i.e., $P(S) = 1$.

$$P(S) = \sum_{i=1}^{\infty} x^2/10^5 = \frac{1}{10^5}\sum_{i=1}^{\infty} x^2 = \infty,$$

$\Rightarrow$ Axiom 1.2 does not hold .

d) Consider the integrand $f(x) = 12x(1-x)^2$. Notice that $f(x)$ is a positive valued function over the domain $S = (0,1)$, implying that any integral of this function over events in the event space will always be nonnegative. Thus Axiom 1.1 holds.

Axiom 1.2 holds, since $P(S) = \int_0^1 12x(1-x)^2\, dx = \left(6x^2 - 8x^3 + 3x^4\right)\Big|_0^1 = 1.$

Finally,

$$P\left[\bigcup_{i\in I} A_i\right] = \int_{x\in\left[\bigcup_{i\in I} A_i\right]} 12x(1-x)^2\, dx = \sum_{i\in I}\int_{x\in A_i} 12x(1-x)^2\, dx = \sum_{i\in I} P(A_i)$$

$\Rightarrow$ Axiom 1.3 holds.

All three axioms hold, implying this is a valid probability set function.

23. a) $S = \{(x,y): x \text{ and } y \in \{1, 2, 3, 4\}\}$

b) $A = \{(x,y): x = 1 \text{ or } 3, y \in \{1, 2, 3, 4\}\}$ (First disk is a 1 or 3).

$B = \{(x,y): y = 1 \text{ or } 2, x \in \{1, 2, 3, 4\}\}$ (Second disk is a 1 or 2).

$N(S) = 16$, $N(A) = 8$, $N(B) = 8$, $N(A \cap B) = 4$.

$$P(A \cap B) = \frac{4}{16} = \left(\frac{8}{16}\right)\left(\frac{8}{16}\right) = P(A)\,P(B) \Rightarrow \text{independence.}$$

c) $A = \{(x, y): x = 1, y \in \{1, 2, 3, 4\}\}$

$B = \{(x, y): y = 1, x \in \{1, 2, 3, 4\}\}$

$N(A) = 4$, $N(B) = 4$, $N(A1B) = 1$

$$P(A \cap B) = \frac{1}{16} = \left(\frac{4}{16}\right)\left(\frac{4}{16}\right) = P(A)\,P(B) \Rightarrow \text{independence.}$$

d) $A = \{(1,1)\}$

$B = \{(x, y): x \text{ and } y \in \{1, 2\}\}$
(neither 3 or 4 is chosen in the selection process)

$N(A) = 1$, $N(B) = 4$, $N(A \cap B) = 1$

$$P(A \cap B) = \frac{1}{16} \neq \frac{1}{16} \cdot \frac{4}{16} = P(A)\,P(B) \Rightarrow \underline{\text{not}} \text{ independent.}$$

25. a) $P(\bar{A} \mid B) = 1 - P(A \mid B) = .02$.

b) $$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid \bar{B})P(\bar{B})} = \frac{(.98)(.05)}{(.98)(.05) + (.02)(.95)}$$

$\Rightarrow P(B \mid A) = .7206$.

c) From Bayes Rule

$$.95 = \frac{r\,(.05)}{r\,(.05) + (1-r)(.95)} \Rightarrow r = .9972.$$

27. a) $S = \{(x, y): x \text{ and } y \in \{0, 1, 2, 3, 4\}\}$.

b) Yes. Let $f(x, y)$ represent the probability of selling $x$ number of printers and $y$ number of computers, as displayed in the table. Then if $A$ is an event contained in $S$,

$$P(A) = \sum_{(x,y)\in A}\sum f(x, y).$$

c) Let $A = \{(x, y): y = 3 \text{ or } 4, x\in \{0, 1, 2, 3, 4\}\}$ (more than two computers sold)

$$\Rightarrow P(A) = \sum_{(x,y)\in A}\sum f(x, y) = .56.$$

Let $B = \{(x, y): x = 3 \text{ or } 4, y\in \{0, 1, 2, 3, 4\}\}$ (more than two printers sold)

$$\Rightarrow P(B) = \sum_{(x,y)\in B}\sum f(x, y) = .50.$$

d) $P(x > 2 \mid y > 2) = \dfrac{P(x>2, y>2)}{P(y>2)} = \dfrac{.40}{.56} = .7143.$

e) $P(x > 2, y > 2) = .10 + .10 + .05 + .15 = .40.$

f) $f(0, 0) = .03,$

$$P(x=0 * y = 0) = \frac{P(x=0, y=0)}{P(y=0)} = \frac{.03}{.08} = .375.$$

29. a) Total number of invoices = 529.
Number of invoices from sales team $C$ = 129.
Probability that a randomly selected invoice is from sales team $C$ = 129/529 =0.2439.

b) Total number of invoices = 529.
Number of invoices over 180 days old = 72.
Probability that a randomly selected invoice is from sales team $C$ = 72/529 =0.1361.

c) Total number of invoices = 529.
Number of invoices over 180 days old and from sales team $C$= 19.
Probability that a randomly selected invoice is over 180 days old and from sales team $C$=19/529 =0.0354.

Sales team $A$ has the lowest probability (0.1493) of being associated with a randomly selected invoice.

31. a)

$i.\ P(A_i) \geq 0 \forall i.$

$ii.\ P(S) = \sum_{i\in\{1,2,3,4\}} P(A_i) = 1$ is possible when the four events are disjoint.

$iii.\ P\left(\bigcup_{i\in\{1,2,3,4\}} A_i\right) = \sum_{i\in\{1,2,3,4\}} P(A_i)$ is possible when the four events are disjoint.

b)

$i.\ P(A_i) \geq 0 \forall i.$

$ii.\ P(S) = 1$ is possible.

$iii.\ P\left(\bigcup_{i\in I} B_i\right) = \sum_{i\in I} P(B_i)$ is possible when $B_i$ are disjoint events.

Eg. Let $P(A_2) = P(A_3) = P(A_4) = 1/3;\ A_2, A_3, A_4$ are disjoint events;
and $A_1 \subset A_2$, which means $A_1, A_2 - A_1, A_3, A_4$ are disjoint events.

Therefore, the assignment of probabilities is actually possible.

c) $P(A_1 \cup A_2) = 0.7 + 0.6 - 0.1 = 1.2 \Rightarrow P(S) \geq 1.2 \neq 1$.

Therefore, the assignement of probabilities is not possible.

d) $A_1 \subset A_1 \cup A_2 \Rightarrow P(A_1) \leq P(A_1 \cup A_2)$ and $A_2 \subset A_1 \cup A_2 \Rightarrow P(A_2) \leq P(A_1 \cup A_2)$.

Since above inequalities (Theorem 1.3) are not satisfied, the assignement of probabilities is

not possible.

e) $A_1 \subset A_2 \subset A_3 \subset A_4 \Rightarrow P(A_1) \leq P(A_2) \leq P(A_3)$.

Since above inequalities (Theorem 1.3) are not satisfied, the assignement of probabilities is not possible.

f) $P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) = .4 + .3 - .5 = .2 \neq P(\varnothing) = 0$.

Therefore, the assignement of probabilities is not possible.

33. a) Define the events:
A: Invoice is from Team A.
B: Invoice is from Team B.

C: Invoice is from Team C.
D: Invoice is from Team D.
X: Age of invoice is less than 120 days.
Y: Age of invoice is 120-180 days.
Z: Age of invoice is over 180 days.
$\Pr(C \mid C \cup D) = N(C) / N(C \cup D) = 129/154 = 0.45583$.

b) The probability that two randomly selected invoices from the pooled set of invoices, selected sequentially without replacement, are both over 180 days old = 72/529 * 71/528=0.0183.

c) $\Pr(C \mid X) = \Pr(C \cap X) / \Pr(X) = N(C \cap X) / N(X) = 45/279 = 0.16129$.

d)

$$\Pr(A \cup B \mid (X \cup Y)) = \Pr((A \cup B) \cap (X \cup Y)) / \Pr(X \cup Y)$$
$$= N((A \cup B) \cap (X \cup Y)) / N(X \cup Y) = 203/457 = 0.4442.$$

35. a) If the airline does not overbook, and only sells 100 tickets for each of their flights, what is the probability that a given flight will fly full?

$(.995)^{100} = .60577$.

b) Using the sales strategy in a), what is the probability that one or more seats for a given flight will be empty?

$1 - (.995)^{100} = 1 - .60577 = .39423$.

c) For the sales strategy in a), if the airline has 10 flight per day from the Seattle-Tacoma airport, what is the probability that all of the flights will fly full?

$\left((.995)^{100}\right)^{10} = .006654$.

d) For the sales strategy in a), what is the probability that there will be one or more empty seats among the 1000 seats available on the airline's 10 flights from the Seattle-Tacoma airport on a given day?

$1 - \left((.995)^{100}\right)^{10} = .993346$.

37. a) Define the events,
M: Pay by cash
C: Pay by credit card
L: Pay by Layaway Plan
X: Purchase below \$100
Y: Purchase \$100-\$500

Z: Purchase over \$500

$P(M) = .40$

$P(X) = .31$

$P(M \cap X) = .20$

$P(M) * P(X) = .40 * .31 = .124 \neq .20$

Therefore, the two events are <u>not</u> independent.

b) Given that the customer pays cash, what is the probability that the customer spends ≤ \$500?

$P(X \cup Y \mid M) = 0.35 / .40 = .875$.

c) Given that the customer pays by credit card, what is the probability that the customer spends ≤\$500 ?

$P(X \cup Y \mid C) = 0.25 / .45 = .5556$.

d) What is the probability that the customer pays by credit card given that the purchase is ≤\$500?

$P(C \mid X \cup Y) = 0.25 / .66 = .3788$.

e) Given that the customer spends \$100 or more, what is the probability that the customer will not pay by cash?

| | **Method of Payment** | | |
|---|---|---|---|
| **Size of Purchase** | **Cash** | **Credit Card** | **Layaway Plan** |
| <\$100 | .20 | .10 | .01 |
| \$100 to \$500 | .15 | .15 | .05 |
| >\$500 | .05 | .20 | .09 |

$P\left(\bar{M} \mid Y \cup Z\right) = P(C \cup L \mid Y \cup Z) = 0.49 / .69 = .7101$.

39. a)

$P\left(A_1 \cap C\right) = 0.20 * 0.0001 = 0.000020,$

$P\left(A_2 \cap C\right) = 0.35 * 0.0002 = 0.000070,$

$P\left(A_3 \cap C\right) = 0.45 * 0.0005 = 0.000225.$

Plant 3 is most likely to have produced the contaminated spinach.
$P(A_3 \mid C) = P(A_3 \cap C) / P(C) = 0.000225 / 0.000315 = .714286$.

b) Plant 1 is least likely to have produced the contaminated spinach.
$P(A_1 \mid C) = P(A_1 \cap C) / P(C) = 0.000020 / 0.000315 = .063492$

c) $P(A_1 \cup A_2 \mid C) = 1 - P(A_3 \mid C) = 1 - .714286 = .285714$.

41. a) What is the probability that the team that wins the first game of the series will go on to win the World Series?

Probability of winning $= \dfrac{\binom{6}{6}+\binom{6}{5}+\binom{6}{4}+\binom{6}{3}}{2^6} = \dfrac{1+6+15+20}{64} = \dfrac{42}{64} = .65625$.

b) What is the probability that a team that has lost the first three games of the series will win the World Series?

Probability of winning $= \dfrac{1}{2^4} = \dfrac{1}{16} = .0625$.

**Chapter 2 – Student Answer Key – Odd Numbers Only**
**Random Variables, Densities, and**
**Cumulative Distribution Functions**

1. a) Not valid. $\sum_{x=0}^{1} f(x) = .8 \neq 1$.

   b) Valid. Definition 2.7a is satisfied.

   c) Not valid. $\int_0^\infty .6e^{-x/4}\,dx = .24 \neq 1$.

   d) Valid. Definition 2.7b is satisfied.

3. a) $Y = I_{[11.75,\ 12.25]}(X)$ $= 1 \Rightarrow$ full
   $= 0 \Rightarrow$ underfilled or overfilled

   b) $R(Y) = \{0,1\}$.

   c) $P(y = 0) = \int_{-\infty}^{11.75} f(x)\;dx + \int_{12.25}^{\infty} f(x)\,dx = 1.38879 \times 10^{-11}$,

   $P(y = 1) = 1 - P(y = 0) = (1 - 1.38879 \times 10^{-11})$,

   $$\Rightarrow f(y) = \left\{\begin{matrix} 1.38879 \times 10^{-11} \\ 1 - 1.38879 \times 10^{-11} \end{matrix}\right\} \text{ when } y = \left\{\begin{matrix} 0 \\ 1 \end{matrix}\right\},$$
   $$= 0 \qquad\qquad \text{elsewhere}$$
   $f(1) \approx 1$.

   d) The range of $X$ is $(-\infty, \approx)$. We cannot place a negative amount of cola in a bottle, nor can we place an infinite amount of cola in a bottle. However, note that for all practical purposes, $x \in A = (11, 13)$, since $P(x \notin A) \approx 0$.

5. a) Assume Star Enterprises maximizes profit,

   $$\Pi(v,w) = \max_x vq(x) - wx.$$

   From the first order condition $x =$ $x = \frac{1}{4}\left(\frac{v}{w}\right)^2$ which yields $\pi = \frac{v^2}{4w}$, which defines the r.v. we seek.

   $$R(\Pi) = \{8.3\underline{3}, 12.50, 25.00, 33.3\underline{3}, 50.00, 75.00, 100.00, 112.50, 225.00\},$$
   $$\Upsilon = \{A : A \subset R(\Pi)\}.$$

b) Since the output and input prices are independent $h(\pi) = f(v)\, g(w)$, then the values of $h(\pi)$ for given values of $\pi{=}\pi(v,w)$ are given below.

| **π** | **h(π) =** | **f(v)** | **g(w)** | **v** | **w** |
|---|---|---|---|---|---|
| 8.3$\underline{3}$ | .06 | (.2) | (.3) | 10 | 3 |
| 12.50 | .06 | (.2) | (.3) | 10 | 2 |
| 25.00 | .08 | (.2) | (.4) | 10 | 1 |
| 33.3$\underline{3}$ | .15 | (.5) | (.3) | 20 | 3 |
| 50.00 | .15 | (.5) | (.3) | 20 | 2 |
| 75.00 | .09 | (.3) | (.3) | 30 | 3 |
| 100.00 | .20 | (.5) | (.4) | 20 | 1 |
| 112.50 | .09 | (.3) | (.3) | 30 | 2 |
| 225.00 | .12 | (.3) | (.4) | 30 | 1 |

$h(\pi) = 0$ for $\pi \notin R(\pi)$,

$P(A) = \sum_{\pi \in A} h(\pi)$ for $A \in \Upsilon$.

c) $P(\{100, 112.50, 225.00\}) = .41$.

d) Let $A = \{\pi\colon \pi > 0\} = R(\Pi) \Rightarrow P(A) = 1$ of making positive profit. Yes, $A$ is certain since $A = R(\Pi)$.

e) Let $A = \{100, 112.50, 225.00\}$ and $B = \{225.00\}$,

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = .2927.$$

7. a) The event of interest is $A = [0,10]$.

$$P(A) = \int_0^{10} .01\exp(-x/100) I_{[0,\infty)}(x)dx = .01\left[-100\exp(-x/100)\Big|_0^{10}\right] = -1\left[\exp(-1/10) - \exp(0)\right]$$
$$= .0952.$$

b) Let $B$ be the event that the screen functions for at least 50,000 hours, i.e., $B = [50, \infty)$.

$$P(B) = \int_{50}^{\infty} .01\exp(-x/100) I_{[0,\infty)}(x)dx = -1\left[\exp(-x/100)\Big|_{50}^{\infty}\right] = -1\left[0 - \exp(-.5)\right]$$
$$= .6065.$$

Let $C = [100, \infty)$ be the event of functioning at least 50,000 hours after the screen has already functioned 50,000 hours.

Then $P(C \mid B) = \dfrac{P(C \cap B)}{P(B)}$.

$$= \frac{\int_{100}^{\infty} .01 \exp(-x/100) I_{(0,\infty)}(x)dx}{\int_{50}^{\infty} .01 \exp(-x/100) I_{(0,\infty)}(x)dx} = \frac{-1\left[\exp(-x/100)\big|_{100}^{\infty}\right]}{-1\left[\exp(-x/100)\big|_{50}^{\infty}\right]} = \exp(-1)/\exp(-.5) = \exp(-.5)$$
$$= .6065.$$

Notice the probability that the screen functions for at least 50,000 hours is the same as the probability that the screen functions for at least an additional 50,000 hours given that it already functioned 50,000 hours. This is a special "memoryless" property of an exponential density function, of which this is an example. Further discussion of this density function will be taken up in Chapter 4.

9. a) Given the units of measurement, the event we are interested in is $pq < 2$.

$$P(pq < 2) = \int_{.1}^{.3} \int_{0}^{2/p} 5pe^{-pq} I_{[.1,.3]}(p) I_{(0,\infty)}(q)\, dq\, dp$$
$$= \int_{.1}^{.3} \left( 5p \frac{e^{-pq}}{-p} \Big|_{0}^{2/p} \right) dp = \int_{.1}^{.3} 5\left[1 - e^{-2}\right] dp = 5p\left[1 - e^{-2}\right]\Big|_{.1}^{.3} = .8647.$$

b) $$f_P(p) = \int_{0}^{\infty} 5pe^{-pq} I_{[.1,.3]}(p) dq$$
$$= \frac{5pe^{-pq}}{-p}\Big|_{0}^{\infty} I_{[.1,.3]}(p) = 5I_{[.1,.3]}(p)$$

$$P(p > .25) = \int_{.25}^{.3} 5\, dp = .25.$$

c) $$f(q \mid p = .20) = \frac{f(.20, q)}{f_P(.20)} = \frac{5(.20)e^{-.20q} I_{(0,\infty)}(q)}{5} = .20e^{-.20q} I_{(0,\infty)}(q)$$

$$P(q > 5 \mid p = .20) = \int_{5}^{\infty} f(q \mid p = .20) dq$$
$$= \int_{5}^{\infty} .20e^{-.20q} dq$$
$$= -e^{-.20q}\Big|_{5}^{\infty} = .3679.$$

d) $$f(q \mid p=.10)=\frac{f(.10,q)}{f_P(.10)}=\frac{5(.10)e^{-.10q}I_{(0,\infty)}(q)}{5}=.10e^{-.10q}I_{(0,\infty)}(q)$$

$$p(q>5 \mid p=.10)=\int_5^\infty .10e^{-.10q}dq=-e^{-.10q}\Big|_5^\infty=.6065.$$

Yes, this makes economic sense. The lower the price, the higher the probability of selling a given quantity of cable.

11. a) $$P(x+y>2)=\sum_{\substack{x+y>2\\ f(x,y)>0}}\sum f(x,y).$$

From the first row of the table we have 2 events satisfying $x+y>2$, we sum their respective probabilities as .01+.03=.04. Similarly, from the second row, .01 + .01 + .003 = .023; from the third .01+.005+.005+.002=.022; and from the fourth, .005+.004+.003+.002+.001=.015. Then summing the probabilities found in each row yields $P(x+y>2)$ = .04 + .023 + .022 + .015 = .10.

b) $$f_Y(y)=\sum_{x\in R(X)}f(x,y)=\sum_{x=0}^{3}f(x,y)$$
$$=.84I_{\{0\}}(y)+.069I_{\{1\}}(y)+.028I_{\{2\}}(y)+.027I_{\{3\}}(y)+.036I_{\{4\}}(y).$$

Then, the probability that more than two gardeners will be absent is given by $P(y>2)=\sum_{y\in\{3,4\}}f_Y(y)=.063.$

c) The random variables $X$ and $Y$ are independent iff $P(x\in A_1, y\in A_2) = P(x\in A_1)P(y\in A_2)$ for all events $A_1$, $A_2$. This condition does not hold in this case. Consider $A_1 = \{x: x=0\}$, $A_2 = \{y: y=0\}$; then $P(x=0, y=0) = .75$, $P(x=0) = .825$, $P(y=0) = .84$ and $.75 \neq (.825)(.84)$ so that $X$ and $Y$ are not independent.

d) $$f(x \mid y=0)=\frac{f(x,0)}{f_Y(0)}=\frac{.75}{.84}I_{\{0\}}(x)+\frac{.06}{.84}I_{\{1\}}(x)+\frac{.025}{.84}I_{\{2\}}(x)+\frac{.005}{.84}I_{\{3\}}(x)$$
$$=.8929I_{\{0\}}(x)+.0714I_{\{1\}}(x)+.0298I_{\{2\}}(x)+.006I_{\{3\}}(x).$$

Then $P(x=0|y=0) = .8929$.

$$f(x \mid y\ge 1)=\frac{\sum_{y=1}^{4}f(x,y)}{\sum_{y=1}^{4}f_Y(y)}=\frac{\sum_{y=1}^{4}f(x,y)}{.16}=\frac{.075}{.16}I_{\{0\}}(x)+\frac{.053}{.16}I_{\{1\}}(x)+\frac{.022}{.16}I_{\{2\}}(x)+\frac{.01}{.16}I_{\{3\}}(x)$$
$$=.4688I_{\{0\}}(x)+.3313I_{\{1\}}(x)+.1375I_{\{2\}}(x)+.0625I_{\{3\}}(x).$$

Then $P(x=0|y\geq 1) = .4688 \Rightarrow$ lower.

13. a) $f(x,y) = \dfrac{\partial^2 F(x,y)}{\partial x \partial y} = \dfrac{1}{20} e^{-(x+5y)/10} I_{(0,\infty)}(x) I_{(0,\infty)}(y)$.

b) $f_X(x) = \dfrac{1}{20} \int_{-\infty}^{\infty} e^{-(x+5y)/10} I_{(0,\infty)}(x) I_{(0,\infty)}(y) dy = \dfrac{1}{10} e^{-x/10} I_{(0,\infty)}(x)$.

c) $F_X(b) = \int_{-\infty}^{b} f_x(x)\, dx = \left[1 - e^{-b/10}\right] I_{(0,\infty)}(b)$.

(Alternatively, $F_X(b) = F(b,\infty)$ by definition)

15. a) $\sum_{x=1}^{\infty} f(x) = 1 \Rightarrow \alpha \sum_{x=1}^{\infty} (1-\beta)^{x-1} = \alpha + \alpha\left(\dfrac{(1-\beta)}{\beta}\right) = 1 \Rightarrow \alpha = \beta \in (0,1)$.

b) Discrete, see Definition 2.4.

c) $f(1) = \beta(1-\beta)^0 = \beta = .05 \Rightarrow f(x) = .05(.95)^{x-1}\, I_{\{1,2,3,\ldots\}}(x)$.

$P(x=10) = f(10) = .0315$.

d) $F(b) = \sum_{\{x \le b:\, x \in \{1,2,3,\ldots\}\}} f(x) = \sum_{\{x \le b:\, x \in \{1,2,3,\ldots\}\}} (.05)(.95)^{x-1} = [1 - .95^{\text{trunc}(b)}]\, I_{[1,\infty)}(b)$.

$P(x > 10) = 1 - F(10) = .5987$.

e) $P(x > 20 | x > 10) = \dfrac{P(x > 20)}{P(x > 10)} = \dfrac{(.95)^{20}}{(.95)^{10}} = .5987$.

17. a) $f(b) = \dfrac{dF(b)}{db} = \dfrac{1}{6} e^{-b/6} I_{(0,\infty)}(b)$.

$P(x \le 6) = \left(1 - e^{-6/6}\right) = .6321$.

b) <u>**NOTE**</u>: This CDF is not properly defined over the entire real line. The definition should read

$$f(b) = \left(\frac{5}{3}\right)\left(.6 - .6^{trunc(b)+1}\right) I_{(0,\infty)}(b)$$

$$\left\{ \begin{array}{l} f(1) = F(1) = \left(\frac{5}{3}\right)\left(.6 - .6^2\right) \\ f(x) = F(x) - F(x-1) = \frac{5}{3}\left(.6^x - .6^{x+1}\right) \text{ for } x \in \{2,3,...\} \\ \quad = 0 \text{ otherwise} \end{array} \right\},$$

$$\left(x \le 6\right) = \left(\frac{5}{3}\right)\left(.6 - .6^7\right) = .9533.$$

19. Let $S = \{(i,j): i,j \in \{1,2,3,4,5,6\}\}$ be the sample space depicting the number of dots facing up on each die after a roll. The r.v. of interest is

$x = \min(i,j)$ for $(i,j) \in S$ with range $R(X) = \{1,2,3,4,5,6\}$.

Let $B_x = \{(i,j): x = \min(i,j), (i,j) \in S\}$, then the probability density function is

$f(x) = P(B_x) = \dfrac{N(B_x)}{N(S)}$ for $x \in R(X)$ and $f(x) = 0$ for $x \notin R(x)$, where $N(S) = 36$.

For example, when $x$=4 then $B = \{(4,4), (4,5), (4,6), (5,4), (6,4)\}$, and $f(4) = 5/36 = .1389$.

The probability space is $\{R(X), \Upsilon, P(B)\}$, where $\Upsilon = \{A: A \subset R(X)\}$

21. a) $R(X, Y) = \{(x,y): x \text{ and } y \in \{0, 1, 2, 3\}, x + y \le 3\}$

$$f(x,y) = \frac{\binom{5}{x}\binom{4}{y}\binom{11}{3-x-y}}{\binom{20}{3}} \text{ for } (x,y) \in R(X,Y)$$
$$= 0 \text{ otherwise.}$$

Note, $f(x,y)$ is the ratio of the number of different ways to choose committees having $x$ economists and $y$ business majors divided by the total number of different ways to choose a committee of 3.

Probability space is $\{S, \Upsilon, P(A)\}$,
where $S = R(X,Y)$, $\Upsilon = \{A: A \subset S\}$, $P(A) = \sum\sum_{(x,y) \in A} f(x,y)$.

b) Let

$$A = \{(x, y) : x \geq 1 \text{ and } y \geq 1\},$$

$$P(A) = f(1,1) + f(1,2) + f(2,1)$$

$$= \frac{\binom{5}{1}\binom{4}{1}\binom{11}{1} + \binom{5}{1}\binom{4}{2} + \binom{5}{2}\binom{4}{1}}{\binom{20}{3}}$$

$$= \frac{220 + 30 + 40}{1140} = .2544.$$

c) $f(0,0) = \dfrac{\binom{11}{3}}{\binom{20}{3}} = .1447.$

d) No, it is not. The probability space does not differentiate between art and political science majors, and so there is no basis for assigning this probability given the result in a).

e)

$$f_X(x) = \sum_{\substack{y=0 \\ (x,y)\varepsilon R(X,Y)}}^{3} f(x,y)$$

$$= \binom{5}{3} \Big/ \binom{20}{3} = .0088 \text{ if } x = 3$$

$$= \binom{5}{2}\left[\binom{4}{1} + \binom{11}{1}\right] \Big/ \binom{20}{3} = .1316 \text{ if } x = 2$$

$$= \binom{5}{1}\left[\binom{4}{2} + \binom{4}{1}\binom{11}{1} + \binom{11}{2}\right] \Big/ \binom{20}{3} = .4605 \text{ if } x = 1$$

$$= \left[\binom{4}{3} + \binom{4}{2}\binom{11}{1} + \binom{4}{1}\binom{11}{2} + \binom{11}{3}\right] \Big/ \binom{20}{3} = .3991 \text{ if } x = 0$$

$$P(x = 3) = f_X(3) = .0088.$$

f)

$$f(y \mid x = 2) = \frac{f(2,y)}{f_X(2)} \text{ for } (2,y) \in R(X,Y)$$

$$= \frac{f(2,0)}{f_X(2)} = \frac{\binom{5}{2}\binom{11}{1}}{\binom{20}{3}(.1316)} = .7332 \text{ for } y = 0$$

$$= \frac{f(2,1)}{f_X(2)} = \frac{\binom{5}{2}\binom{4}{1}}{\binom{20}{3}(.1316)} = .2667 \text{ for } y = 1.$$

$P(y{=}0|x{=}2) = f(y{=}0|x{=}2) = .7332.$

g)

$$f(y \mid x \geq 2) = \sum_{\substack{x \geq 2 \\ (x,y) \in R(X,Y)}} f(x,y) \Big/ \sum_{x \geq 2} f_X(X)$$

$$= \frac{f(2,0) + f(3,0)}{.1404} = \frac{\binom{5}{2}\binom{11}{1} + \binom{5}{3}}{\binom{20}{3}(.1404)} = .7497 \text{ for } y = 0$$

$$= \frac{f(2,1)}{.1404} = \frac{\binom{5}{2}\binom{4}{1}}{\binom{20}{3}(.1404)} = .2499 \text{ for } y = 1.$$

(Does not add to 1 because of rounding.)

$P(y{=}0|x{\geq}2) = .7497.$

h) No. Since the conditional probabilities in f) and g) depend on which event for $x$ we are conditioning on, $X$ and $Y$ cannot be independent.

23. a)

$$f_3(x_3) = \sum_{x_1=0}^{3} \sum_{x_2=0}^{3} f(x_1, x_2, x_3) = \left(\frac{.48}{1 + x_3}\right) I_{\{0,1,2,3\}}(x_3),$$

$P(x_3 = 3) = f_3(3) = .12.$

b)

$$f_{12}(x_1,x_2)=\sum_{x_3=0}^{3} f(x_1,x_2,x_3)=\left(\frac{25}{12}\right)(.004)(3+2x_1+x_2)\prod_{i=1}^{2} I_{\{0,1,2,3\}}(x_i)$$
$$=\frac{1}{120}(3+2x_1+x_2)\prod_{i=1}^{2} I_{\{0,1,2,3\}}(x_i).$$

$$P(x_1>1,x_2>1)=\sum_{x_1=2}^{3}\sum_{x_2=2}^{3} f_{12}(x_1,x_2)$$
$$=\frac{1}{120}[12\ +\ 20\ +\ 10]=.35.$$

c)

$$f(x_1 \mid x_2\ge 2)=\frac{\sum_{x_2=2}^{3} f_{12}(x_1,x_2)}{\sum_{x_2=2}^{3} f_2(x_2)},$$

where $f_2(x_2)=\sum_{x_1=0}^{3} f_{12}(x_1,x_2)=\frac{1}{120}(24+4x_2)I_{\{0,1,2,3\}}(x_2)$
$$=(.2+x_2/30)I_{\{0,1,2,3\}}(x_2).$$

Thus, $f(x_1 \mid x_2\ge 2)=\dfrac{\frac{1}{120}[6+4x_1+5]I_{\{0,1,2,3\}}(x_1)}{\left(\frac{68}{120}\right)}$
$$=\frac{1}{68}[11+4x_1]I_{\{0,1,2,3\}}(x_1),$$
$$P(x_1\le 1 \mid x_2\ge 2)=\sum_{x_1=0}^{1} f(x_1 \mid x_2\ge 2)=.3824.$$

d) No, ($X_1$, $X_2$, $X_3$) are dependent r.v.s since the conditional density for $X_1$, given an event for $X_2$, depends on the event for $X_2$.

Yes, ($X_1$, $X_2$) is independent of $X_3$ since

$$f(x_1,x_2,x_3)=f_{12}(x_1,x_2)f_3\left(x_3=\frac{1}{120}(3+2x_1+x_2)\right)\left(\frac{.48}{1\ +\ x_3}\right)\prod_{i=1}^{3} I_{\{0,1,2,3\}}(x_i)$$
$$=\frac{(.004)(3+2x_1+x_2)}{(1\ +\ x_3)}\prod_{i=1}^{3} I_{\{0,1,2,3\}}(x_i).$$

e)

$$f(x_1, x_2 \mid x_3 = 0) = \frac{f(x_1, x_2, 0)}{f_3(0)}$$

$$= \frac{f_{12}(x_1, x_2) f_3(0)}{f_3(0)} \text{ by independence of } (X_1, X_2) \text{ and } X_3$$

$$= f_{12}(x_1, x_2),$$

$$P(x_1 > 1, x_2 > 1 \mid x_3 = 0) = P(x > 1, x_2 > 1) = .35 \text{ (from b)).}$$

f) $\Pi = 20X_1 + 30X_2 + 60X_3 - 150$.

$R(\Pi) = \{\pi: \pi = 20x_1 + 30x_2 + 60x_3 - 150, x_i \in \{0,1,2,3\}, i=1,2,3\}$.

Let $A(\pi) = \{(x_1, x_2, x_3): \pi = 20x_1 + 30x_2 + 60x_3 - 150, x_i \in \{1, 2, 3\}, i=1,2,3\}$.

$$f(\pi) = \sum\sum\sum_{(x_1,x_2,x_3)\in A(\pi)} f(x_1, x_2, x_3).$$

$$P(\pi > 0) = \sum_{\substack{\pi>0 \\ \pi\in R(\Pi)}} f(\pi) = \sum\sum\sum_{(x_1,x_2,x_3)\in B} f(x_1, x_2, x_3) = .364.$$

where $B$ is the set of $(x_1, x_2, x_3)$ points, $x_i \in \{0,1,2,3\}$ $\forall i$, for which $\pi > 0$.

25. The marginal densities are

$$f_1(x_1) = \sum_{x_2=0}^{1}\sum_{x_3=0}^{1} f(x_1, x_2, x_3) = \frac{1}{2} I_{\{0,1\}}(x_1),$$

$$f_2(x_2) = \sum_{x_1=0}^{1}\sum_{x_3=0}^{1} f(x_1, x_2, x_3) = \frac{1}{2} I_{\{0,1\}}(x_2),$$

$$f_3(x_3) = \sum_{x_1=0}^{1}\sum_{x_2=0}^{1} f(x_1, x_2, x_3) = \frac{1}{2} I_{\{0,1\}}(x_3),$$

$$f_{12}(x_1, x_2) = \sum_{x_3=0}^{1} f(x_1, x_2, x_3) = \frac{1}{4} I_{\{0,1\}}(x_2) I_{\{0,1\}}(x_3),$$

$$f_{13}(x_1, x_3) = \sum_{x_2=0}^{1} f(x_1, x_2, x_3) = \frac{1}{4} I_{\{0,1\}}(x_1) I_{\{0,1\}}(x_3),$$

$$f_{23}(x_2, x_3) = \sum_{x_1=0}^{1} f(x_1, x_2, x_3) = \frac{1}{4} I_{\{0,1\}}(x_2) I_{\{0,1\}}(x_3).$$

a) For $i,j \in \{1,2,3\}$ and $i \neq j$ $\quad f_{ij}(x_i, x_j) = f_i(x_i) f_j(x_j)$.

$\Rightarrow$ pairwise independent.

b) Note for $x_1 = x_2 = x_3 = 0$.

$f(0,0,0) = .20 \neq f_1(0)\, f_2(0)\, f_3(0) = .125.$

$\Rightarrow$ jointly dependent.

27. a) Not a PDF since $\int_0^\infty f(x)dx = 10 \neq 1$.

b) Not a PDF since $\sum_{x=0,1,2...} f(x) = 1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + .... = \frac{4}{3} \neq 1$.

c) Not a PDF since $\int_0^\infty f(x)dx = \frac{1}{\ln\left(\frac{1}{4}\right)} = .72135 \neq 1$.

d) Can be a PDF since $\int_0^1 \int_0^1 f(x,y)dxdy = 1$ and $f(x,y) \geq 0$ for $\forall x,y$.

29. a) $f(c) = \frac{e^c}{\left(1+e^c\right)^2}$ for $c \in (-\infty, \infty)$.

b) $f(c) = \begin{cases} 2x^{-3} \text{ for } c \in (1,\infty) \\ 0 \text{ otherwise} \end{cases}$.

c) $f(c) = \begin{cases} (.5)^c \text{ for } c \in \{1,2,3,...\} \\ 0 \text{ otherwise} \end{cases}$.

d) $f(c_1,c_2) = 6c_1^2 c_2 I_{[0,1]}(c_1) I_{[0,1]}(c_2)$.

e) $f(c_1,c_2) = e^{-c_1} e^{-c_2} I_{[0,\infty)}(c_1) I_{[0,\infty)}(c_2)$.

31. a) $\Pr(x = 5; n = 10) = f(5) = 252(.02)^5(.98)^5 = 7.2892 \times 10^{-7}$.

b)

$$f(0) = 1(.02)^0(.98)^{10} = 0.817072807$$
$$f(1) = 10(.02)^1(.98)^9 = 0.166749552$$
$$f(2) = 45(.02)^2(.98)^8 = 0.015313734$$
$$f(3) = 120(.02)^3(.98)^7 = 0.000833401$$
$$f(4) = 210(.02)^4(.98)^6 = 2.9764\times10^{-5}$$
$$f(5) = 252(.02)^5(.98)^5 = 7.2892\times10^{-7}$$
$$\Pr(x \le 5; n=10) = f(0)+f(1)+f(2)+f(3)+f(4)+f(5) = 0.999999987$$

c) $\Pr(x = 0; n = 10) = f(0) = 1(.02)^0(.98)^{10} = 0.817072807$.

d) $\Pr(x = 0) \mid f(x \le 5) = 0.817072807 / 0.999999987 = 0.817072817$.

33. a) $F(z) = \begin{cases} 1-(.5)^{\text{floor}(z+1)}, \text{for } z \ge 0 \\ 0 \quad \text{otherwise} \end{cases}$ where $\text{floor}(z) \equiv$ round down the value $z$.

b) $\Pr(z < 10) = F(9) = 1-(.5)^{10} = 0.9990234$.

c) $\Pr(z > 3) = 1 - F(3) = (.5)^4 = 0.0625$.

d) $\Pr(z = 0 \mid z \le 2) = F(0)/F(2) = (1-.5)/(1-(.5)^3) = 0.5714286$.

35. a) $\Pr(t = w) = \sum_{t=0}^{4} \frac{3t}{100} = 0.3$.

b) $f_t(t) = \begin{cases} t(t+1)/40, \text{for } t \in \{0,1,2,3,4\} \\ 0 \quad \text{elsewhere} \end{cases} \Rightarrow \Pr(t \le 2) = f_t(0) + f_t(1) + f_t(2) = 0.20$.

c) $f_w(w) = \begin{cases} (5-w)(2+w)/50, \text{for } w \in \{0,1,2,3,4\} \\ 0 \quad \text{elsewhere} \end{cases} \Rightarrow \Pr(w \ge 3) = f_w(3) + f_w(4) = 0.32$.

d) $T$ and $W$ are not independent random variables since $f(t,w) \ne f_t(t) f_w(w)$.

37. a)

$$f(x,y) = e^{-(x+y)} I_{[0,\infty)}(x) I_{[0,\infty)}(y) = e^{-x} I_{[0,\infty)}(x) e^{-y} I_{[0,\infty)}(y) = f_x(x) f_y(y)$$

$\Rightarrow x$ and $y$ are independent.

b)

$$f(x,y)=\frac{x(1+y)}{300}I_{\{1,2,3,4,5\}}(x)I_{\{1,2,3,4,5\}}(y)=\frac{x}{15}I_{\{1,2,3,4,5\}}(x)\frac{(1+y)}{20}I_{\{1,2,3,4,5\}}(y)=f_x(x)f_y(y)$$

$\Rightarrow x$ and $y$ are independent.

c)

$$\begin{aligned} f(x,y,z) &= 8xyz\ I_{[0,1]}(x)I_{[0,1]}(y)I_{[0,1]}(z) \\ &= 4xy\ I_{[0,1]}(x)I_{[0,1]}(y)2z\ I_{[0,1]}(z)=f_{xy}(x,y)f_z(z) \\ &= 4xz\ I_{[0,1]}(x)I_{[0,1]}(z)2y\ I_{[0,1]}(y)=f_{xz}(x,z)f_y(y) \\ &= 2x\ I_{[0,1]}(x)4yz\ I_{[0,1]}(y)I_{[0,1]}(z)=f_x(x)f_{yz}(y,z) \\ &= 2x\ I_{[0,1]}(x)2y\ I_{[0,1]}(y)2z\ I_{[0,1]}(z)=f_x(x)f_y(y)f_z(z) \end{aligned}$$

$\Rightarrow x, y$ and $z$ are independent.

d)

$$\begin{aligned} f(x_1,x_2,x_3) &= \frac{.5^{x_1}.2^{x_2}.75^{x_3}}{10}\prod_{i=1}^{3}I_{\{0,1,2,\ldots\}}(x_i) \\ &= \frac{.5^{x_1}.2^{x_2}}{2.5}I_{\{0,1,2,\ldots\}}(x_1)I_{\{0,1,2,\ldots\}}(x_2)\frac{.75^{x_3}}{4}I_{\{0,1,2,\ldots\}}(x_3)=f_{12}(x_1,x_2)f_3(x_3) \\ &= \frac{.5^{x_1}}{2}I_{\{0,1,2,\ldots\}}(x_1)\frac{.2^{x_2}.75^{x_3}}{5}I_{\{0,1,2,\ldots\}}(x_2)I_{\{0,1,2,\ldots\}}(x_3)=f_1(x_1)f_{23}(x_2,x_3) \\ &= \frac{.5^{x_1}.75^{x_3}}{8}I_{\{0,1,2,\ldots\}}(x_1)I_{\{0,1,2,\ldots\}}(x_3)\frac{.2^{x_2}}{1.25}I_{\{0,1,2,\ldots\}}(x_2)=f_{13}(x_1,x_3)f_2(x_2) \\ &= \frac{.5^{x_1}}{2}I_{\{0,1,2,\ldots\}}(x_1)\frac{.2^{x_2}}{1.25}I_{\{0,1,2,\ldots\}}(x_2)\frac{.75^{x_3}}{4}I_{\{0,1,2,\ldots\}}(x_3)=f_1(x_1)f_2(x_2)f_3(x_3) \end{aligned}$$

$\Rightarrow x_1, x_2$ and $x_3$ are independent.

39. a)

$$f(x)=\begin{cases}\frac{1}{36}I_{\{1,9,16,25,36\}}(x)+\frac{1}{18}I_{\{2,3,5,8,10,15,18,20,24,30\}}(x)+\frac{1}{12}I_{\{4\}}(x)+\frac{1}{9}I_{\{6,12\}}(x) & \text{for } x\in R(X), \\ 0 & \text{elsewhere}\end{cases}$$

$R(X)=\{1,2,3,4,5,6,8,9,10,12,15,16,18,20,24,25,30,36\}$.

b) $\Pr(X\geq 16)=\dfrac{11}{36}$

c)

$$f(x,y)=\begin{cases} \frac{1}{36}I_{\{1\}}(x)I_{\{2\}}(y)+\frac{1}{36}I_{\{4\}}(x)I_{\{4\}}(y)+\frac{1}{36}I_{\{9\}}(x)I_{\{6\}}(y)+\frac{1}{36}I_{\{16\}}(x)I_{\{8\}}(y)+\frac{1}{36}I_{\{25\}}(x)I_{\{10\}}(y) \\ \quad +\frac{1}{36}I_{\{36\}}(x)I_{\{12\}}(y)+\frac{1}{18}I_{\{2\}}(x)I_{\{3\}}(y)+\frac{1}{18}I_{\{3\}}(x)I_{\{4\}}(y)+\frac{1}{18}I_{\{4,6\}}(x)I_{\{5\}}(y) \\ \quad +\frac{1}{18}I_{\{5,8\}}(x)I_{\{6\}}(y)+\frac{1}{18}I_{\{6,10,12\}}(x)I_{\{7\}}(y)+\frac{1}{18}I_{\{12,15\}}(x)I_{\{8\}}(y)+\frac{1}{18}I_{\{18,20\}}(x)I_{\{9\}}(y) \\ \quad +\frac{1}{18}I_{\{24\}}(x)I_{\{10\}}(y)+\frac{1}{18}I_{\{30\}}(x)I_{\{11\}}(y) \qquad \text{for } (x,y)\in R(X,Y) \\ 0 \quad \text{elsewhere} \end{cases}$$

$R(X,Y)=\{(1,2),(2,3),(3,4),(4,4),(4,5),(5,6),(6,5),(6,7),(8,6),(9,6),(10,7),(12,7),(12,8),(15,8),$ $(16,8),(18,9),(20,9),(24,10),(25,10),(30,11),(36,12)\}$.

d) $\Pr((X\geq 16)\,\&\,(Y\geq 8))=\frac{11}{36}$.

e) No. Because $f(x,y)\neq f_x(x)f_y(y)$.

f) $f(x\mid y=7)=\begin{cases}\frac{1}{3}I_{\{6,10,12\}}(x) & \text{for } (x)\in R(X)=\{6,10,12\} \\ 0 & \text{elsewhere}\end{cases}$.

g) $\Pr(X\geq 10\mid Y=7)=\frac{2}{3}$.

**Chapter 3 – Student Answer Key – Odd Numbers Only**
**Mathematical Expectation and Moments**

1. a) Probability of warranty claim = $P(x \leq 3)$ for each TV = $\int_0^3 .005\, e^{-.005} dx = .015$.

Let $z_i = 1 \Rightarrow$ warranty claim for $i^{\text{th}}$ TV
$= 0 \Rightarrow$ no claim.

$P(z_i = 1) = .015$, $P(z_i = 0) = .985$, $\mathrm{E}Z_i = 1(.015) + 0(.985) = .015$,
$\mathrm{E}\sum_{i=1}^{100} Z_i = \sum_{i=1}^{100} \mathrm{E}Z_i = 100(.015) = 1.5$.

b) $\mathrm{E}X = \int_0^{\infty} x f(x)\, dx = 200$ years.

3. a) Notice that

$$f_Y(y) = \sum_{x=0}^{3} f(x,y) = .39 I_{\{0\}}(y) + .285 I_{\{1\}}(y) + .155 I_{\{2\}}(y) + .10 I_3(y) + .07 I_{\{4\}}(y),$$

$$\Rightarrow \mathrm{E}Y = \sum_{y=0}^{4} y f_Y(y) = 1.175,$$

and

$$f_X(x) = \sum_{y=0}^{4} f(x,y) = .505 I_{\{0\}}(x) + .265 I_{\{1\}}(x) + .12 I_{\{2\}}(x) + .11 I_{\{3\}}(x),$$

$$\Rightarrow \mathrm{E}X = \sum_{x=0}^{3} x f_X(x) = .835.$$

Then letting $G = 200X + 100Y$ be Al's weekly commission, we have

$$\begin{aligned}\mathrm{E}G &= \mathrm{E}(200X + 100Y) = 200\mathrm{E}X + 100\mathrm{E}Y,\\ &= 200(.835) + 100(1.175) = \$284.50.\end{aligned}$$

Total weekly pay is $T = 100 + G$, so that

$$\mathrm{E}T = \mathrm{E}(100 + G) = 100 + \mathrm{E}G = 100 + 284.50 = \$384.50.$$

b) Letting $C = 100Y$ be the weekly commission from selling compact cars, we have

$\mathrm{E}C = \mathrm{E}(100Y) = 100\mathrm{E}Y = 100(1.175) = \$117.50$.

Letting $L = 200X$ be the commission from selling luxury cars, we have

$EL = 200EX = 200(.835) = \$167.$

c) The conditional expectation is given as

$$E(200X \mid y=4) = \sum_{x=0}^{3}\left(200x\right) f(x \mid y=4) = \sum_{x=0}^{3}\left(200x\right)\frac{f(x,4)}{f_Y(4)}$$
$$= 200\left[(0)(.03/.07) + (1)(.02/.07) + (2)(.01/.07) + (3)(.01/.07)\right]$$
$$= \$200.$$

d) Letting $N = (1-.38)(100 + G)$ be weekly pay net of taxes, we have

$$EN = E(.62(100+G)) = 62 + .62EG$$
$$= 62 + .62(284.50) = \$238.39.$$

5. a)

$$E(Q \mid p) = \int_{-\infty}^{\infty} q\,\frac{f(p,q)}{f_P(p)}\,dq$$
$$= \int_{-\infty}^{\infty} q\left[\frac{2pe^{-pq} I_{[1/2,1]}(p) I_{(0,\infty)}(q)}{2I_{[1/2,1]}(p)}\right] dq$$
$$= \frac{1}{p},\ p \in \left[\frac{1}{2},1\right].$$

b) Graph $E(Q \mid p) = 1/p$, for $p \in \left[\frac{1}{2},1\right]$.

c) $E(Q \mid p = .75) = 1/.75 = 1.3\underline{3}$ million gallons

d) Total revenue $TR = p \cdot q$.

$$E(TR) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(p \cdot q\right)\left(2pe^{-pq} I_{[1/2,1]}(p) I_{(0,\infty)}(q)\right) dq\,dp$$
$$= 1 \text{ million dollars.}$$

7. a)

$$EX = \int_0^1 x3x^2 dx = .75,$$
$$EX^2 = \int_0^1 x^2 3x^2 dx = .6,$$
$$\operatorname{var}\left(X\right) = EX^2 - \left(EX\right)^2 = .0375.$$

$$\int_0^{\text{med}(X)} 3x^2 dx = 1/2 \Rightarrow \text{med}(X) = .7937,$$

$$\mu_3 = \int_0^1 (x - .75)^3 3x^2 dx = -.00625 \neq 0 \Rightarrow \text{ not symmetric.}$$

b)

$$\text{E}X = 2,$$

$$\text{E}X^2 = \int_1^\infty x^2 2x^{-3} dx = 2\ln x \,|_1^\infty = \infty \Rightarrow \text{var}(X) \text{ does not exist,}$$

$$\int_1^{\text{med}(X)} 2x^{-3} dx = 1/2 \Rightarrow \text{med}(X) = 1.414.$$

$f(x)$ maximized at $x$=1⇒mode ($X$) = 1.

$\mu_3$ does not exist since $\sigma_X^2$ does not exist. Density is skewed right (examine graph).

c) E$X$ undefined, since $\int_0^\infty f(x)dx = \infty$ and $\int_{-\infty}^0 f(x)dx = -\infty$ .

⇒var ($X$) does not exist.

$$\int_{-\infty}^{\text{med}(X)} \frac{1}{\pi}\left(\frac{1}{x^2+1}\right) dx = 1/2 \quad \Rightarrow \text{med}(X) = 0,$$

$f(x)$ maximized at $x = 0$ ⇒ mode($X$) = 0,

$\mu_3$ does not exist. Density is symmetric (examine graph).

d) E$X$ = .8,
E$X^2$ = 1.28,
var($X$) = E$X^2$−(E$X$)$^2$ = .64,
$P(x \geq \text{med}(X)) \geq 1/2$, $P(x \leq \text{med}(X)) \geq 1/2 \Rightarrow \text{med}(X) = 1$,
$f(x)$ is maximized at either $x$=0 or $x$=1, where $f(x)$ = .4096, so both are modes.
$\mu_3$ = .384≠0, so density is not symmetric.

9. Letting $X$ = MPG attained by purchasers of the line of pickups, we are given $\mu$ = E$X$ = 17 and $\sigma = \sqrt{\text{E}(X-\mu)^2} = .25$ .

The event in question is

$$B = [x : 16 < x < 18] = [x : \mu - 4\sigma < x < \mu + 4\sigma] = [x : |x - \mu| < 4\sigma].$$

Applying Chebyshev's inequality (Corollary 3.4), we have

$$P(B) = P[|x - \mu| < 4\sigma] \geq 1 - \frac{1}{4^2} = .9375.$$

The lower bound on the probability of this event is .9375. Notice we are able to state this lower bound in the absence of any information regarding the algebraic form of the density of $X$.

11. a)

$$\mathrm{E}X_1 = \left.\frac{dM_X(t)}{dt_1}\right|_{t=0} = \frac{d}{dt_1}\exp\left[\left(.07t_1 + .11t_2\right) + \frac{1}{2}\left(.225\times10^{-3}t_1^2 - .6\times10^{-3}t_1t_2 + .625\times10^{-3}t_2^2\right)\right]\Bigg|_{t=0}$$

$$= \left[.07 + .225\times10^{-3}t_1 - .3\times10^{-3}t_2\right]M_X(t)|_{t=0} = .07.$$

Similarly, $\mathrm{E}X_2 = .11$.

b)

$$\operatorname{cov}(X_1, X_2) = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1X_2} \\ \sigma_{X_2X_1} & \sigma_{X_2}^2 \end{bmatrix} = \begin{bmatrix} .225\times10^{-3}\ x\,10^{-3} & -.3\times10^{-3} \\ -.3\times10^{-3} & .625\times10^{-3} \end{bmatrix}.$$

For example, $\sigma_{x_1x_2} = \mathrm{E}(X_1X_2) - (\mathrm{E}X_1)(\mathrm{E}X_2) = \left.\frac{d^2M_X(t)}{dt_1dt_2}\right|_{t=0} - \left(\left.\frac{dM_X(t)}{dt_1}\right|_{t=0}\right)\left(\left.\frac{dM_X(t)}{dt_2}\right|_{t=0}\right)$

which yields $\sigma_{X_1X_2}^2 = .0074 - (.07)(.11) = -.3 \times 10^{-3}$.

c) $\operatorname{corr}(X_1, X_2) = \begin{bmatrix} 1 & \rho_{X_1X_2} \\ \rho_{X_2X_1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -.80 \\ -.80 & 1 \end{bmatrix}$. Since $|\rho_{X_1X_2}| \neq 1$, the linear relationship $x_1 = \alpha_1 + \alpha_2 x_2$ does not hold.

d) Let $a_i$ be the amount invested in the $i^{\text{th}}$ project, $i$=1,2.

Expected return $\mathrm{E}Y = a_1\mathrm{E}X_1 + a_2\mathrm{E}X_2 = a_1(.07) + a_2(.11)$.

Because project 2 has higher returns, choose $a_2 = 1000$ and $a_1 = 0$.

$\operatorname{var}(Y) = \operatorname{var}(1000\,X_2) = (1000)^2\operatorname{var}(X_2) = 625$.

e) Objective is to $\min\limits_{a_1, a_2} \operatorname{var}(a_1X_1 + a_2X_2)$ s.t. $a_1 + a_2 = 1000$.

Let $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, then $L = a'\,\Phi\,a + \lambda(1000 - a_1 - a_2)$.

First order conditions are

$$\left\{\begin{array}{l} \dfrac{\partial L}{\partial a} = 2\Phi a - \lambda\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \\ \dfrac{\partial L}{\partial \lambda} = 1000 - \alpha_1 - \alpha_2 = 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{lc} .45\times 10^{-3} a_1 - .6\times 10^{-3} a_2 - \lambda & = 0 \\ -.6\times 10^{-3} a_1 + 1.25\times 10^{-3} a_2 - \lambda & = 0 \\ a_1 + a_2 & = 1000 \end{array}\right\},$$

$$\Rightarrow \begin{bmatrix} a_1 \\ a_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 637.93 \\ 362.07 \\ .0698 \end{bmatrix}.$$

E$Y = a_1$ E$X_1 + a_2$ E$X_2 = 84.48$.

f) E$U(M)$ = E$(5M)$ = 5E$M$ and as in part d), choose $a_1 = 0$, $a_2 = 1000$.
var($Y$) = var(1000 $X_2$) = 625.

g) E$U(M)$ = E $5M^2$ = 5 E$M^2$ = 5(var($M$) + (E$M$)$^2$).

$$\max_{a_1, a_2} 5(a'\Phi\ a + (a'u)^2)\ s.t.\ a_1 + a_2 = 1000.$$

Upon substituting the constraint into the objective function, we obtain

$$\max_{a_1} g(a_1) = .025625 a_1^2 + .074 a_1 (1000 - a_1) + .063625 (1000 - a_1)^2.$$

Notice, $\dfrac{dg(a_1)}{da_1} < 0 \quad \forall a_1 \in [0, 1000]$.

Hence the optimal investment is still $a_1 = 0$, $a_2 = 1000$. The variance of this investment is var(1000 $X_2$) = 625.

h) Because expected utility increases both with respect to increased mean returns and increased variance of returns (i.e., the individual is risk seeking) project 2 is chosen in both f) and g) because of its higher returns and higher risk.

13. a) The marginal density of $x_1$ is

$$f(x_1) = \int_0^1 \int_0^1 \frac{1}{3}(x_1 + 2x_2 + 3x_3) I_{(0,1)}(x_1)\, dx_2\, dx_3$$
$$= \frac{1}{3}(x_1 + 5/2) I_{(0,1)}(x_1)$$

and hence,

$$\text{E}X_1 = \int_{-\infty}^{\infty} x_1 \cdot \frac{1}{3}(x_1 + 5/2) I_{(0,1)}(x_1)\, dx_1 = .5278.$$

Likewise,

$$\mathrm{E}X_2 = \int_{-\infty}^{\infty} x_2 \cdot \frac{1}{3}\left(2x_2+2\right) I_{(0,1)}\left(x_2\right) dx_2 = .5556.$$

$$\mathrm{E}X_3 = \int_{-\infty}^{\infty} x_3 \cdot \frac{1}{3}\left(3x_3+3/2\right) I_{(0,1)}\left(x_3\right) dx_3 = .5833.$$

b) $$\mathrm{E}\left(\frac{X_1+X_2+X_3}{3}\right) = \frac{1}{3}\left(\mathrm{E}X_1+\mathrm{E}X_2+\mathrm{E}X_3\right) = \frac{1}{3}(.5278+.5556+.5833) = .5556.$$

c) The conditional density is

$$f\left(x_1,x_2 \mid x_3=.90\right) = \frac{f\left(x_1 x_2,.90\right)}{f_3(.90)} = \frac{\frac{1}{3}\left(x_1+2x_2+3(.90)\right) I_{(0,1)}\left(x_1\right) I_{(0,1)}\left(x_2\right)}{\frac{1}{3}(3/2\ +\ 3(.90))}$$
$$= \frac{\left(x_1+2x_2+2.7\right) I_{(0,1)}\left(x_1\right) I_{(0,1)}\left(x_2\right)}{4.2}$$

and marginal conditional densities are

$$f_1\left(x_1 \mid x_3=.90\right) = \int_{-\infty}^{\infty}\left(\frac{x_1+2x_2+2.7}{4.2}\right) I_{(0,1)}\left(x_1\right) I_{(0,1)}\left(x_2\right) dx_2 = \frac{x_1+3.7}{4.2} I_{(0,1)}\left(x_1\right),$$
$$f_2\left(x_2 \mid x_3=.90\right) = \frac{2x_2+3.2}{4.2} I_{(0,1)}\left(x_2\right).$$

Using the above information

$$\mathrm{E}\left(x_1 \mid x_3=.90\right) = \int_0^1 x_1\left(\frac{x_1+3.7}{4.2}\right) dx_1 = .5198,$$
$$\mathrm{E}\left(x_2 \mid x_3=.90\right) = \int_0^1 x_2\left(\frac{2x_2+3.2}{4.2}\right) dx_2 = .5397.$$

d) Amount of sewage treated is

$Y = 100{,}000(X_1 + X_2) + 250{,}000(X_3)$.

Then, $\mathrm{E}Y = 100{,}000(\mathrm{E}X_1 + \mathrm{E}X_2) + 250{,}000(\mathrm{E}X_3) = 254{,}165$ gallons.

15. a) $P(x = 7 \text{ or } 11) = 6/36 + 2/36 = 8/36$ and $P(x \neq 7 \text{ or } 11) = 1-8/36 = 28/36$, then

$\mathrm{E}X = 2Z(8/36) - Z(28/36) = -\frac{1}{3}Z \neq 0$. Hence, the game is not fair.

b) Accounting for the cost of each spin the net payoffs are

| $x$ | $f(x)$ |
|---|---|
| $-.40$ | $1/3$ |
| $.20$ | $1/6$ |
| $.20$ | $1/6$ |
| $1.40$ | $1/12$ |
| $-.20$ | $1/4$ |

The $\mathrm{E}(X) = \Sigma x f(x) = 0,$ which implies a fair game.

c) Since $\mathrm{E}X = \sum_{j=1}^{\infty} 2^j \left(\frac{1}{2}\right)^j \to \infty$, then the player could never be charged enough to make the game fair.

17. a) The conditional expectation of quantity sold as a function of price is given by

$$\mathrm{E}(Q \mid p) = \int_{-\infty}^{\infty} q \frac{f(p,q)}{f_P(p)} dq .$$

Since,

$$f_P(p) = \int_{-\infty}^{\infty} .5pe^{-pq} I_{[3,5]}(p) I_{(0,\infty)}(q) dq$$
$$= .5pI_{[3,5]}(p) \int_0^{\infty} e^{-pq} dq = .5I_{[3,5]}(p),$$

we have

$$\mathrm{E}(Q \mid p) = \int_0^{\infty} pqe^{-pq} dq .$$

Treating $p$ as unspecified constant, and integrating by parts (i.e., $\int v\, du = uv - \int u\, dv$, where $v = pq \Rightarrow dv = p\, dq$, $du = e^{-pq}\, dq \Rightarrow u = -(1/p)\, e^{-pq}$ we have

$$\mathrm{E}(Q \mid p) = -qe^{-pq} \Big|_0^{\infty} - \int_0^{\infty} -e^{-pq} dq = \lim_{q \to \infty} \left( \frac{1}{-pe^{pq}} \right) + \frac{1}{p} \text{ (L'Hospital's rule)}$$
$$= \frac{1}{p} \text{ for } p \in [3,5].$$

b) The graph of the regression curve is the graph of the function $y = 1/p$ for $p \in [3, 5]$.

$\mathrm{E}(Q \mid p = 3.50) = .2857,$
$\mathrm{E}(Q \mid p = 4.50) = .2222.$

19. a) Markov's inequality states that for nonnegative valued $g(S)$,

$P[g(s)\geq a] \leq \mathrm{E}g(S)/a \qquad \forall\ a>0.$

In this application $g(S) = S$ and $a = 30$, so that

$P[s\geq 30]\leq 20/30 = .67.$

With only the knowledge that E$S$ = 20, we can place an upper bound of .67 on the probability of the event in question. The store manager is incorrect.

b) $\mathrm{var}(S) = \sigma^2 = 1.96 \Rightarrow \sigma = \sqrt{1.96}$.

Then $P[s\epsilon(10,30)]$

$$= P\left[20 - \frac{10}{\sqrt{1.96}}\sigma < s < 20 + \frac{10}{\sqrt{1.96}}\sigma\right]$$
$$= P\left[\frac{-10}{\sqrt{1.96}}\sigma < s - 20 < \frac{10}{\sqrt{1.96}}\sigma\right]$$
$$= P\left[|s-20| < \frac{10}{\sqrt{1.96}}\sigma\right] \geq 1 - 1/\left(10/\sqrt{1.96}\right)^2 = .9804 \quad \text{by Chebyshev's inequality.}$$

21. a) $$\mathrm{E}Y = \int_{-\infty}^{\infty} y\left(y^{-2}\right)I_{(1,\infty)}\left(y\right)dy = \int_{1}^{\infty}\frac{1}{y}dy = \lim_{b\to\infty}\ln\left(y\right)\Big|_1^b = \lim_{b\to\infty}\ln\left(b\right) = \infty$$

$\Rightarrow$E$Y$ does not exist.

b) By the contrapositive of Theorem 3.23, E$Y^s$ for $s > 1$ also does not exist.

23. a)

$$f_X(x) = \int_{-\infty}^{\infty}\frac{2}{5}(3x+2y)I_{[0,1]}(x)I_{[0,1]}(y)dy$$
$$= \frac{2}{5}I_{[0,1]}(x)\int_0^1\left(3x+2y\right)dy = \frac{2}{5}(3x+1)I_{[0,1]}(x).$$

Then, $P(x>.75) = \int_{.75}^{\infty}\frac{2}{5}(3x+1)I_{[0,1]}(x) = \frac{2}{5}\left(\frac{2}{3}x^2 + x\right)\Big|_{t=0} = .3625.$

b) The regression curve of $Y$ on $X$ is defined as

$$\mathrm{E}(Y \mid x) = \int_{-\infty}^{\infty} y \frac{f(x,y)}{f_X(x)} dy$$
$$= \int_{-\infty}^{\infty} y \frac{(3x+2y)}{(3x+1)} I_{[0,1]}(y) dy \quad \text{(using } f_X(x) \text{ from part a)}$$
$$= \frac{1}{3x+1} \int_0^1 \left(3xy + 2y^2\right) dy = \frac{9x+4}{18x+6} \text{ for } x \in [0,1].$$

Then we have

$$\mathrm{E}(Y \mid x = .75) = \frac{9(.75)+4}{18(.75)+6} = .5513.$$

Since E($Y|x$) depends on $x$, $Y$ and $X$ are not independent.

c) Letting $R$ = 1.25($X$·1000) + 1.40($Y$·1000) = 1250$X$ + 1400$Y$ be weekly revenue, we have

$$\mathrm{E}(R \mid x = .75) = \mathrm{E}[(1250X + 1400Y) \mid x = .75]$$
$$= 1250\,\mathrm{E}(X \mid x = .75) + 1400\mathrm{E}(Y \mid x = .75)$$
$$= 1250(.75) + 1400(.5513) = \$1709.32.$$

25. a) To maximize expected utility ($\alpha$ = 0)

E$U(\Pi)$ = E$P \cdot q - C(q)$, where

$$\mathrm{E}p = \int_0^5 p\left(.048\left(5p - p^2\right)\right) dp = .048\left[\frac{5}{3}p^3 - \frac{1}{4}p^4\right]\Bigg|_0^5 = 2.5,$$

So

$$\max_q \mathrm{E}\,U(\Pi) = \max_q 2.5q - \frac{1}{2}q^2 \Rightarrow \frac{\partial \mathrm{E}U(\Pi)}{\partial q} = 2.5 - q = 0 \text{ or } q = 2.5.$$

b) var($\Pi$) = E($\Pi^2$)−(E$\Pi$)$^2$, which requires E($\Pi^2$) = E($Pq$−1/2 $q^2$)$^2$ = (E$P^2$)$q^2$−(E$P$) $q^3$ + 1/4$q^4$ and E$P^2$ = $\int_0^5 p^2 \cdot (.048(5p-p^2))dp$ = 7.5. Hence, var($\Pi$) = (7.5 $q^2$−2.5$q^3$ + 1/4 $q^4$)−(2.5$q$ −1/2$q^2$)$^2$ = 1.25 $q^2$.

So,

$\max_q$ (2.5$q$−1/2$q^2$)−$\alpha$(1.25$q^2$) ⇒ $\dfrac{\partial \mathrm{E}U(\pi)}{\partial q}$ = 2.5−(1 + 2.5$\alpha$)$q$ = 0 or $q = \dfrac{2.5}{1+2.5\alpha}$. For $q$ = 1 ⇒ $\alpha$ = .60.

c) For $\alpha$=1 $\Rightarrow q = .7143$. If prices are given, $\max_q \; pq - \frac{1}{2}q^2$, which yields optimal $q = p$. So, $p$=.7143 would induce the firm to the desired level of output.

27. a) Letting $f(x,y)$ represent the bivariate density function as given in the table and $\Pi$ = (22,000 − 20,000)$X$ + (7500−6500)$Y$ = 2000X + 1000Y represent daily profit above dealer cost, we have (from Theorem 3.8)

$$\mathrm{E}\Pi = \sum_{y=0}^{3}\sum_{x=0}^{3}(2000x + 1000y)\, f(x,y) = 2000\sum_{x=0}^{3} x f_X(x) \; + \; 1000\sum_{y=0}^{3} y f_Y(y),$$
$$= 2000\ \mathrm{E}X + 1000\ \mathrm{E}Y.$$

Since $f_X(x) = .11\, I_{\{0\}}(x) + .40 I_{\{1\}}(x) + .35\, I_{\{2\}}(x) + .14\, I_{\{3\}}(x) \Rightarrow \mathrm{E}X = \sum_{x=0}^{3} x f_X(x) = 1.52$, and $f_Y(y) = .14 I_{\{0\}}(y) + .24 I_{\{1\}}(y) + .36 I_{\{2\}}(y) + .26 I_{\{3\}}(y)$ ,

$\Rightarrow \mathrm{E}Y = \sum_{y=0}^{3} y\, f_Y(y) = 1.74$ we have $\mathrm{E}\Pi = 2000(1.52) + 1000(1.74) = \$4780$.

b) Total daily profit can be presented as

$T = 2000X + 1000Y - 4000$.

The following table lists range elements of $T$ along with associated probabilities in parenthesis:

| No. of Mini-Rovers Sold ($Y$) \ No. of land Yachts Sold ($X$) | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | -4,000 (.05) | -2,000 (.05) | 0 (.02) | 2,000 (.02) |
| 1 | -3,000 (.03) | -1,000 (.10) | 1,000 (.08) | 3,000 (.03) |
| 2 | -2,000 (.02) | 0 (.15) | 2,000 (.15) | 4,000 (.04) |
| 3 | -1,000 (.01) | 1,000 (.10) | 3,000 (.10) | 5,000 (.05) |

$t$

$(f(t))$

Letting $A = \{t: t > 0\}$, then $P(A) = \sum_{t\in A} f(t) = .57$.

c)

$$\mathrm{E}(Y \mid x=0)=\sum_{y=0}^{3} y \frac{f(0,y)}{f_X(0)}=(0)\left(\frac{.05}{.11}\right)+(1)\left(\frac{.03}{.11}\right)+(2)\left(\frac{.02}{.11}\right)+(3)\left(\frac{.01}{.11}\right)=.91,$$

$$\mathrm{E}(Y \mid x=2)=\sum_{y=0}^{3} y \frac{f(2,y)}{f_X(2)}=(0)\left(\frac{.02}{.35}\right)+(1)\left(\frac{.08}{.35}\right)+(2)\left(\frac{.15}{.35}\right)+(3)\left(\frac{.10}{.35}\right)=1.94.$$

Since $E(Y|x)$ depends on $x$, $Y$ and $X$ are not independent.

29. a) If fertilizer is applied at the rate of 27 lbs/ac, i.e., $x$=27, then $Y = 30e^{\varepsilon}$. Notice that $Y$ is montonically increasing in ε so that $y > 50$ requires that $\varepsilon > 1n(5/3)$. Then at $x$=27, $P(y > 50)$ = P(ε > 1n(5/3)), and $P(\varepsilon > 1n(5/3)) = \int_{\ln(5/3)}^{\infty} 3\, e^{-3\varepsilon} I_{(0,\infty)}(\varepsilon)\, d\varepsilon = -e^{-3\varepsilon} \Big|_{\ln(5/3)}^{\infty} = .216.$

b)

$$\mathrm{E}Y = \int_{-\infty}^{\infty} \left(10x^{1/2}e^{\varepsilon}\right)\left(3e^{-3\varepsilon}\right) I_{(0,\infty)}(\varepsilon)\, d\varepsilon = 30x^{1/3}\int_{0}^{\infty} e^{-2\varepsilon} d\varepsilon \text{ (note, } x \text{ is not random)}$$

$$= 30x^{1/3}\left[-\frac{1}{2}e^{-2\varepsilon}\Big|_{0}^{\infty}\right] = 15x^{1/3}.$$

Then if we let $R = 3Y - 0.20x$ be return over fertilizer cost, we have

E$R$ = E($3Y-0.20x$) = 3E$Y-0.20x$ = $3(15x^{1/3})-0.20x = 45x^{1/3}-0.20x$,

which, at $x$=27, is equal to \$129.60.

c) In part b) we found E$Y$ and since

$$\mathrm{E}Y^2 = \int_{-\infty}^{\infty} \left(100x^{2/3}e^{2\varepsilon}\right)\left(3e^{-3\varepsilon}\right) I_{(0,\infty)}(\varepsilon) d\varepsilon = 300x^{2/3}\int_{0}^{\infty} e^{-\varepsilon} d\varepsilon = 300x^{2/3},$$

it follows that

var($Y$) = $300x^{2/3}-(15x^{1/3})^2 = 75x^{2/3}$ (recall var($Y$) = E$Y^2-(\mathrm{E}Y)^2$).

When $x$=27, var($Y$) = 675. Yes, the variance of $Y$ changes with different values of $x$.

31. a) $\rho_{x_1x_2} = \dfrac{\sigma_{X_1X_2}}{\sigma_{X_1}\sigma_{X_2}} = \dfrac{.0001}{\sqrt{.00011}\sqrt{.0001}} = .9535$

b) Use Theorem 3.37 to plot line of best prediction, and the discussion following the theorem can be used to motivate the degree of fit.

c) $P(|x_1 - x_2| < .01)$
$= P(|x_1 - x_2 + \mu - \mu| < .01)$ $(EX_1 = EX_2 = \mu)$
$= P(|Y - \mu| < .01)$ (note $Y = X_1 - X_2 + \mu$ and $EY = \mu$)
$\geq 1 - \frac{1}{k^2}$ (Chebyshev's Inequality).

where $k\sigma = .01$. Since

$$\sigma^2 = \text{var}(Y) = \text{var}\left(X_1 - X_2 + \mu\right) = [1\ -1]\begin{bmatrix} .00011 & .0001 \\ .0001 & .0001 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = .00001,$$

then

$$k = \frac{.01}{\sqrt{.00001}} = 3.1623 \text{ and } 1 - \frac{1}{k^2} = .90 \Rightarrow P(|x_1 - x_2| < .01) \geq .90.$$

Yes, agree.

33. Find $P(3 < t < 9)$ without knowing the probability density function of $T$. Apply Chebyshev's Inequality

$$P(3 < t < 9) = P(|t - 6| < 3) \geq 1 - \frac{1}{(2.4496)^2} = .8333,$$

where $3 = k\sigma = k(1.2247)$ or $k = 2.4496$.

35. a) Expected profit =\$ (50−25−10,000(1−.998)) = \$ 5.

b) \$ (25+10,000(1−.998)) = \$ 45.

c) \$ 10,000(.998−.995)= \$ 30.

d) If the premium is fixed the company maximizes its profit by selling the entire portfolio of policies to the 25 year old population.

37. a) $f(0) = \frac{4}{50}; f(1) = \frac{5}{50}; f(2) = \frac{8}{50}; f(3) = \frac{13}{50}; f(4) = \frac{20}{50}.$
Mean = 2.8
Median = Any number in the interval [2,3]
Mode=4
.10 quantile= Any number in the interval [1,2]
.90 quantile= Any number in the interval [4,5]

b) Mean = $\int_0^\infty .5xe^{-x/2} = 2$

Median $= -2\ln(.5) = 1.386294$

Mode= 0

.10 quantile= $-2\ln(.1) = 4.605170$

.90 quantile= $-2\ln(.9) = 0.210721$

c) Mean = $\int_0^1 3x^3 = 0.75$

Median = $(.5)^{1/3} = .793701$

Mode=1

.10 quantile= $(.1)^{1/3} = .464159$

.90 quantile= $(.9)^{1/3} = .965489$

d) Mean = $\sum_{x=1}^{\infty}\left(.05x(.95)^{x-1}\right) = 20$

Median = Any number in the interval [13,14] since $\dfrac{\ln(1-.5)}{\ln(.95)} = 13.513$

Mode= 1

.10 quantile= Any number in the interval [2,3] since $\dfrac{\ln(1-.1)}{\ln(.95)} = 2.054$

.90 quantile= Any number in the interval [44,45] since $\dfrac{\ln(1-.9)}{\ln(.95)} = 44.891$

39. a) No. $Y = 25 + 2R - .05R^2 = 25 + 2R - .05\left[(R-10)^2 + 20R - 100\right] = 30 + R - .05(R-10)^2$

$$\mathrm{E}(Y) = 30 + \mathrm{E}(R) - .05\mathrm{E}(R-10)^2 = 30 + 10 - .05\mathrm{E}(R-10)^2 \le 40$$

b) $\mathrm{E}(Y) = 25 + 2\mathrm{E}(R) - .05\mathrm{E}\left(R^2\right) \le 25 + 2(15) - .05(15)^2 = 43.75$

c)

$$\mathrm{E}(Y) = 25 + 2\mathrm{E}(R) - .05\mathrm{E}\left(R^2\right) = 25 + 2\mathrm{E}(R) - .05\left(\mathrm{var}(R) + \left(\mathrm{E}(R)\right)^2\right)$$

$$= 25 + 2(15) - .05\left(5 + (15)^2\right) = 43.5$$

41. a) Using Markov's inequality, the upper bound of Pr($D$>40)=E($D$)/40=0.5.

b) Using Chebyshev's inequality, $\left(20 - 4\sqrt{5}\right) \le D \le \left(20 - 4\sqrt{5}\right) \Rightarrow D \in (11.05573, 28.944272)$.

43. a) $\left.\dfrac{\partial M_{(P,Q)}(\mathbf{t})}{\partial t_1}\right|_{\mathbf{t}=\mathbf{0}}$ = \$ 125.

b) $\left.\dfrac{\partial M_{(P,Q)}(\mathbf{t})}{\partial t_2}\right|_{\mathbf{t}=\mathbf{0}}$ = 500 kegs.

c) $\left.\dfrac{\partial M_{(P,Q)}(\mathbf{t})}{\partial t_1 \partial t_2}\right|_{\mathbf{t}=\mathbf{0}}$ = (125 x 500) − 100 = \$62,400.

45. a)

$$\mathrm{E}(U(\pi)) = \mathrm{E}(\pi) = q\mathrm{E}(P) - c(q),$$

$$\mathrm{E}(P) = \int_0^\infty pf(p)dp = \int_0^\infty .5pe^{-.5p}dp = 2$$

$\Rightarrow \mathrm{E}(U(\pi)) = 2q - \left(.5q + .1q^2\right) = 1.5q - .1q^2$ is maximized when $q = 7.5$.

b)

$$\mathrm{E}(U(\pi)) = \mathrm{E}(\pi) - \alpha\left[\mathrm{var}(\pi)\right] = \mathrm{E}(\pi) - \alpha\left[q^2 \ \mathrm{var}(P)\right] = q\mathrm{E}(P) - c(q) - \alpha\left[q^2 \ \mathrm{var}(P)\right],$$

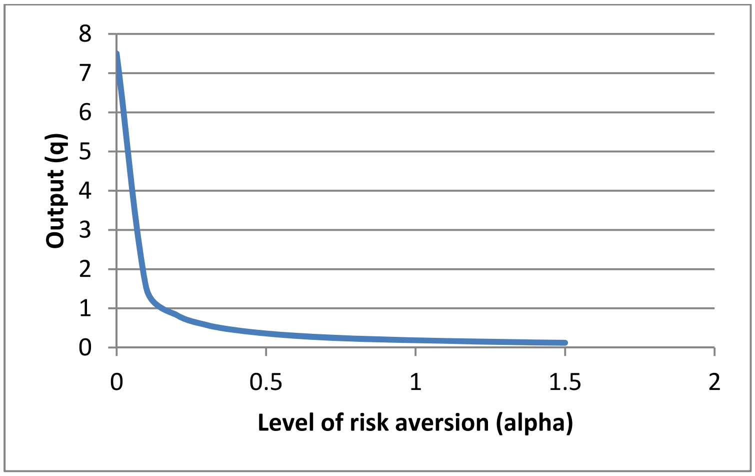
$$\mathrm{E}(P^2) = \int_0^\infty p^2 f(p)dp = \int_0^\infty .5p^2 e^{-.5p}dp = 8,$$

$$\mathrm{var}(P) = \mathrm{E}(P^2) - \left[\mathrm{E}(P)\right]^2 = 8 - 4 = 4$$

$$\Rightarrow \mathrm{E}(U(\pi)) = 2q - \left(.5q + .1q^2\right) - 4\alpha q^2 = 1.5q - (.1 + 4\alpha)q^2.$$

This is maximized when $q = \dfrac{.75}{\left(.1 + 4\alpha\right)}$.

When $q = 5, \alpha = .0125$.




c) When $a = .1,\ q = \dfrac{.75}{(.1 + 4\alpha)} = 1.5.$

When no price uncertainty: $\mathrm{E}\left(U(\pi)\right) = Pq - \left(.5q + 1q^2\right)$ is maximized when $q = 5P - 2.5$ for $q = 1.5, P = .8$.

47. a)

Return, $R = aX$ where $a = \begin{bmatrix} 5000 & 5000 \end{bmatrix}$,
$\mathrm{E}(R) = a\mathrm{E}(X) = \$1{,}250$,
$\mathrm{Var}(R) = a\mathrm{Cov}(X)a' = \$110{,}250$.

b) The investor should invest his entire portfolio in $X_1$.

Let, $R = aX$ where $a = \begin{bmatrix} \theta & (10000 - \theta) \end{bmatrix}$,

$\mathrm{E}(R) = a\mathrm{E}(X) = 500 + .15\theta$ is maximized when $\theta$=10,000.

c) The investor should invest his entire portfolio in $X_2$.

Let, $R = aX$ where $a = \begin{bmatrix} \theta & (10000 - \theta) \end{bmatrix}$,

$\mathrm{Var}(R) = a\mathrm{Cov}(X)a' = .0361\theta^2 + 38\theta + 10000$ is minimized when $\theta$=0.

d) The investor should invest his entire portfolio in $X_2$.

Let, $R = aX$ where $a = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} = \begin{bmatrix} \theta & (10000 - \theta) \end{bmatrix}$,

$\mathrm{E}(R) = a\mathrm{E}(X) = 500 + .15\theta,$

$\mathrm{Var}(R) = a\mathrm{Cov}(X)a' = .0361\theta^2 + 38\theta + 10000,$

$\mathrm{E}(U(R)) = \mathrm{E}(R) - .01\ \mathrm{Var}(R) = 400 - .23\theta - .000361\theta^2$ is maximized when $\theta$=0.

49. a)

$$\hat{Y} = a + bX,$$

$$a = \mathrm{E}(Y) - \left(\frac{\sigma_{xy}}{\sigma_x^2}\right)\mathrm{E}(X) \text{ and } b = \frac{\sigma_{xy}}{\sigma_x^2} \Rightarrow a = 1 \text{ and } b = \frac{2}{5}.$$

b) $$(\rho_{xy})^2 = \left(\frac{2}{\sqrt{2}\sqrt{5}}\right)^2 = (.6325)^2 = .4.$$

# Chapter 4 – Student Answer Key – Odd Numbers Only
## Parametric Families of Density Functions

1. False. Since random sampling <u>without</u> replacement, we utilize the hypergeometric density not the binomial density as given. The probability is given by

$$P(x \le 3) = \sum_{x=0}^{3} \frac{\binom{k}{x}\binom{100-k}{20-x}}{\binom{100}{20}}.$$

3. a)

$$f(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,2,\ldots,n\}}(x) = \exp\left[\ln\binom{n}{x} + \ln\left(\frac{p}{1-p}\right)^x + \ln(1-p)^n\right] I_{\{0,1,2,\ldots,n\}}(x)$$

$$= \exp\left[\ln\binom{n}{x} + x\ln\left(\frac{p}{1-p}\right) + \ln(1-p)^n\right] I_{\{0,1,2,\ldots,n\}}(x).$$

From Definition 4.3, let $k$=1 and fix $n$, then

$c(p) = \ln\left(\frac{p}{1-p}\right), g(x) = x,\ d(p) = \ln(1-p)^n,\ z(x) = \ln\binom{n}{x}$, and $A$ = {0,1,2,...$n$}. It follows that the binomial family, for a fixed value of $n$, is a member of the exponential class of densities.

b) Let $k$=1, then $c(\lambda) = \ln\lambda$, $g(x) = x$, $d(\lambda) = -\lambda$, $z(x) = -\ln(x!)$, and $A$={0,1,2,...,}.

c) Let $k$=1 and fix $r$, then $c(p) = \ln(1-p)$, $g(x)=x$, $d(p)=\ln\left(\frac{p}{1-p}\right)^r$, $z(x) = \ln\binom{x-1}{r-1}$, and $A$={$x$: $x$=$r$, $r$+1, $r$+2,...}.

d) Let $k$=$m$ and fix $n$, then
$c_i(p) = \ln(p_i)$ and $g_i(x) = x_i$ for $i$=1,...,$m$

$d(p) = 0$, $z(x) = \ln\left(\frac{n!}{\prod_{i=1}^m x_i!}\right)$, and $A = \left\{x_1,\ldots,x_m \mid x_i = 0,1,2,\ldots,n \forall i, \sum_{i=1}^m x_i = n\right\}$.

e) Let $k$=2, then

$c_1(\Theta) = \alpha$, $c_2(\Theta) = \beta$, $g_1(x) = \ln(x)$, $g_2(x) = \ln(1-x)$, $d(\Theta) = \ln\left(\frac{1}{B(\alpha,\beta)}\right)$, $z(x) = -\ln(x(1-x))$,

and $A$ = (0,1).

5. Proof: Recall the geometric density is $f(x;p)=\begin{cases} p(1-p)^{x-1}, & \text{for } x=1,2,3,\ldots \\ 0 & \text{otherwise} \end{cases}$,

$$P(x>s+t\,|\,x>s)=\frac{\sum_{x=s+t+1}^{\infty} p(1-p)^{x-1}}{\sum_{x=s+1}^{\infty} p(1-p)^{x-1}}=\frac{\sum_{x=s+t+1}^{\infty}(1-p)^{x}}{\sum_{x=s+1}^{\infty}(1-p)^{x}}=\frac{\frac{(1-p)^{s+t+1}}{p}}{\frac{(1-p)^{s+1}}{p}}=(1-p)^{t},$$

$$\left(\text{Hint: Geometric series } \sum_{x=i}^{\infty} r^{x}=\frac{r^{i}}{(1-r)} \text{ for } |r|<1\right).$$

and

$$p(x>t)=\sum_{x=t+1}^{\infty} p(1-p)^{x-1}=\sum_{x=t+1}^{\infty}(1-p)^{x}\left(\frac{p}{1-p}\right)=\left(\frac{p}{1-p}\right)\left(\frac{(1-p)^{t+1}}{p}\right)=(1-p)^{t}$$
$$\Rightarrow P(x>s+t\,|\,x>s)=P(x>t).$$

7. a) $E[V] = 0$ and $\text{var}(V) = 16 \times 10^6$.

b) If $p = 4$, then $E[Q] = 50{,}000$ and

$$P(q>50{,}000)=P\left(\frac{q-50{,}000}{4{,}000}>0\right)=P(z>0)=\frac{1}{2}.$$

This assumes $Q \sim N(50{,}000,\ 16 \times 10^6)$.

c) If $p = 4.50$, then $Q \sim N(43{,}750,\ 16 \times 10^6)$ and

$$P(q>50{,}000)=P\left(z>\frac{50{,}000-43{,}750}{4{,}000}\right)=p(z>1.5625)=.059.$$

d) Find $p$ such that

$$P(q>50{,}000)=P\left(z>\frac{50{,}000-E[Q]}{4{,}000}\right)=P(z>-1.645)=.95$$
$$\Rightarrow \frac{50{,}000-E[Q]}{4{,}000}=-1.645$$
$$\Rightarrow p=3.47.$$

e) No. Assuming $V$ could actually be normally distributed would imply that $Q$ could be negative for a given value of $p$. Yet, in reality $Q$ is never negative. The approximation may be acceptable if $P(Q < 0) \approx 0$, which it is.

9. Refer to Table B.3.

a) $P(y > 27.488) = .025$.

b) $P(6.262 < y < 27.488) = P(y > 6.262) - P(y > 27.488) = .975 - .025 = .95$.

c) $c = 24.996$.

d) From Corollary 4.2, we know that $Z \sim \chi^2_{10}$. Thus $c = 18.307$.

11. Refer to Theorems 4.10 and 4.11.

a) The regression curve is $\mathrm{E}(X_1 \mid x_2) = 5 - \frac{1}{3}(x_2 - 8)$ and $\mathrm{E}(X_1 \mid x_2 = 9) = 5 - \frac{1}{3}(9-8) = 4.67$.

b) $\sigma^2(X_1|x_2 = 9) = (2) - (-1)(1/3)(-1) = 1.67$.

c) $P(x_1 > 5) = P(z > 0) = 1/2, Z = \dfrac{X_1 - 5}{\sqrt{2}} \sim N(0,1)$.

Using the conditional mean and variance from above,

$$P(x_1 > 5 \mid x_2 = 9) = P\left(z > \frac{5 - 4.67}{\sqrt{1.67}}\right) = P(z > .2554) = 1 - F(.2554) \approx .40,$$

(linearly interpolating in Table B.1).

13. The geometric density is appropriate in this instance.

a) Letting $X$ = the number of attempts necessary to get the first light, the probability space is

$\{R(X), \Upsilon, P(A)\}$,

where $R(X) = \{1, 2, 3, ...\}$, $\Upsilon = \{A: A \subset R(X)\}$, $P(A) = \sum_{x \in A} .95(.05)^{x-1}$.

b) $P(x > 5) = 1 - P(x \leq 5) = 1 - \sum_{x=1}^{5} .95\,(.05)^{x-1} = 3.125 \times 10^{-7}$.

c) For the geometric density, we know that $\mu = 1/p$, where $p$ = the probability of obtaining a Type A outcome (e.g., successfully lighting) on each of the independent Bernoulli trials.

In this case, $\mathrm{E}X = \mu = 1/.95 = 1.05$.

d) Letting $R$ be the value of the award, we have $R$ = 1,000,000 $Z$ where $Z = I_{\{6,7,8,...\}}(X)$ is an r.v. indicating whether or not the customer is paid the \$1,000,000. Then E$R$ = 1,000,000 E$Z$ and, from Theorem 3.3,

E$Z = P(X > 5) = 3.125 \times 10^{-7}$,

so that

E$R$ = 1,000,000 $(3.125 \times 10^{-7})$ = .3125.

15. The exponential density is appropriate here.

a) Letting $X$ = the number of hours the display functions before failure, the probability space is

$\{R(X), \Upsilon, P(A)\}$

where $R(X) = (0, \infty)$, $\Upsilon = \{A: A$ is a Borel set in $R(X)\}$, $P(A) = \int_{x \in A} \frac{1}{30{,}000} e^{-x/30{,}000} I_{(0,\infty)}(x)$.

b) $P(x \geq 20{,}000) = \int_{20{,}000}^{\infty} \frac{1}{30{,}000} e^{-x/30{,}000}\, dx = -e^{-x/30{,}000} \Big|_{20{,}000}^{\infty} = .5134.$

c) By "memoryless property," $P(x > s + t | x > s) = P(x > t)\ \forall t$ and $s > 0$, so that $P(x > 30{,}000 | x > 10{,}000) = P(x > 20{,}000) = .5134$ (from b)).

d) No, it is not more likely. Given the memoryless property, the watch is essentially as good as new while functioning.

17. The beta density is an appropriate choice (as a continuous approximation in this case) for modeling experiments whose outcomes are in the form of proportions:

$R(X) = [0,1]$, $\Upsilon \equiv$ {all Borel sets of $R(X)$}, $P(A) = \int_{x \in A} \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1} I_{(0,1)}(x)\, dx.$

a) Since $\mu = .40$, $\sigma^2 = .04$ one can solve

$.4 = \alpha/(\alpha+\beta)$,
$.04 = (\alpha\beta)/[(\alpha+\beta)^2(\alpha+\beta+1)]$.

simultaneously for $\alpha$, $\beta$. Thus, $\alpha = 2$, $\beta = 3$, and $P(x > .50) = \int_{.5}^{1} 12x(1-x)^2 dx = .3125$.

b) $P(x < .25) = \int_{0}^{.25} 12x\ (1-x)^2 dx = .2617.$

c) The median is derived by solving $\int_0^{med(x)} 12x\ (1-x)^2\, dx = 1/2 \Rightarrow \text{med}(x) = .3857$.
The mode occurs at $x = 1/3$, where $f(1/3;\, 2,3)$ is a maximum.

19. Since sampling without replacement characterizes this experiment, and since there are three categories of accounts, the multivariate hypergeometric distribution is a useful choice (see p. 188).

$$f\left(x_1, x_2, x_3; 200, 140, 45, 15, 5\right) = \frac{\binom{140}{x_1}\binom{45}{x_2}\binom{15}{x_3}}{\binom{200}{5}}; x_i \in \{0,1,\ldots,5\}, \sum_{i=1}^{3} x_i = 5 .$$

a) It is easiest in this case to use the (marginal) hypergeometric distribution for $X_3$ (recall p. 188).

$$P\left(x_3 = 0\right) = \frac{\binom{15}{0}\binom{185}{5}}{\binom{200}{5}} = .6744 .$$

b) $P(x_3 \leq 1) = P(x_3 = 0) + P(x_3 = 1) = .9538$.

c) Using the multivariate hypergeometric,

$$P\left(x_1 = 3, x_2 = 1, x_3 = 1\right) = \frac{\binom{140}{3}\binom{45}{1}\binom{15}{1}}{\binom{200}{5}} = .1191 .$$

d) Using the marginal hypergeometric densities of $X_1$, $X_2$, $X_3$

$$\mathrm{E}\left(X_1\right) = \frac{nk_1}{M} = 3.5, \mathrm{E}\left(X_2\right) = 1.125, \mathrm{E}\left(X_3\right) = .375 .$$

21.

$$f(x;p)=\begin{cases} p(1-p)^{x-1} & x\in\{1,2,3,...\} \\ 0 & \text{otherwise} \end{cases},$$

$$P(x>s+t\mid x>s)=\frac{\sum_{x+t+1}^{\infty}p(1-p)^{x-1}}{\sum_{s+1}^{\infty}p(1-p)^{x-1}}=\frac{(1-p)^{s+t}}{(1-p)^{s}}=(1-p)^{t}$$

$$P(x>t)=\sum_{t+1}^{\infty}p(1-p)^{x-1}=(1-p)^{t}$$

$$\Rightarrow P(x>s+t\mid x>s)\equiv P(x>t),\forall(s,t)\in\{1,2,3,...\}\times\{s,s+1,s+2,...\}.$$

23. a) Since the inspection is without replacement we can model using a hypergeometric density function. An appropriate probability model assuming the number of defectives is $x$ is:

$$\{R(X),\Upsilon,P(A)\},R(X)=\{0,1,2\},\Upsilon=\{A:A\subset R(X)\},P(A)=\sum_{x\in A}\frac{\binom{2}{x}\binom{98}{10-x}}{\binom{100}{10}}$$

b)

$$P(x=0)=\frac{\binom{2}{0}\binom{98}{10}}{\binom{100}{10}}=\frac{90(89)}{100(99)}=.809091.$$

c) $\mathrm{E}(X)=\dfrac{nk}{M}=\dfrac{(10)(2)}{(100)}=.2$ where $n$ is the number of draws, $k$ is the number of defectives and $M$ is the total number of disk players.

d) Now,

$$P(x=0)=\frac{\binom{2}{0}\binom{98}{20}}{\binom{100}{20}}=\frac{80(79)}{100(99)}=.638384.$$

25. a) We can model the given problem using a multivariate hypergeometric distribution as follows.

$$f(x_1,x_2,x_3;100,50,35,15,5)=\frac{\binom{50}{x_1}\binom{35}{x_2}\binom{15}{x_3}}{\binom{100}{5}},x_i\in\{0,1,2,3,4,5\},\sum_{i=1}^{3}x_i=5.$$

Here, $x_1$ denotes being current, $x_2$ being past due and $x_3$ being delinquent.

$$f(x_3;100,85,15,5)=\frac{\binom{85}{5-x_3}\binom{15}{x_3}}{\binom{100}{5}},x_3\in\{0,1,2,3,4,5\},$$

$$P(x_3=0)=\frac{\binom{85}{5}\binom{15}{0}}{\binom{100}{5}}=\frac{(85)(84)(83)(82)(81)}{(100)(99)(98)(97)(96)}=.435683.$$

b)

$$P(x_3\le 1)=P(x_3=0)+P(x_3=1)=\frac{\binom{85}{5}\binom{15}{0}}{\binom{100}{5}}+\frac{\binom{85}{4}\binom{15}{1}}{\binom{100}{5}}$$

$$=\frac{(85)(84)(83)(82)\left[(81)+(15)(5)\right]}{(100)(99)(98)(97)(96)}=.839094.$$

c)

$$f(x_1=3,x_2=1,x_3=1)=\frac{\binom{50}{3}\binom{35}{1}\binom{15}{1}}{\binom{100}{5}}=.136676.$$

d) $\mathrm{E}(X_i)=\dfrac{nk_i}{M}\Rightarrow \mathrm{E}(X_1)=2.5,\mathrm{E}(X_2)=1.75,\mathrm{E}(X_3)=.75.$

27. a) Since the number of occasions a player can role the die before moving a game piece is not finite, we can define a geometric distribution to answer the given questions. Each role of die can be treated as an independent Bernoulli process and the probability of rolling either a one or a six is $1/3$. An appropriate probability model assuming the number of roles to move one game piece is $x$ is:

$$\{R(X),\Upsilon,P(A)\}, R(X)=\{1,2,3...\}, \Upsilon=\{A: A\subset R(X)\}, P(A)=\sum_{x\in A} p(1-p)^{x-1} = \sum_{x\in A}\frac{1}{3}\left(\frac{2}{3}\right)^{x-1}.$$

b) For a geometric distribution $F(b)=\left[1-(1-p)^{trunc(b)}\right]I_{[1,\infty)}(b)$

$$\Rightarrow 1-F(2)=\left(\frac{2}{3}\right)^2=\frac{4}{9}=.444444.$$

c) $\mathrm{E}(X)=\dfrac{1}{p}=3$.

d) Using the result $P(x>s+t\,|\,x>s)=P(x>t)$ we can show that the expected number of rolls to move the $(j+1)^{th}$ game piece after moving the $j^{th}$ game piece is 3. Since the events are independent and sequential the expected number of dies to move all four game pieces $=4(3)=12.$

It is also possible to answer this part using a negative binomial model. Assuming the number of roles to move all four game pieces is $x$:

$$\{R(X),\Upsilon,P(A)\}, R(X)=\{4,5,6,...\}, \Upsilon=\{A: A\subset R(X)\},$$

$$P(A)=\sum_{x\in A}\frac{(x-1)!}{3!(x-4)!}p^4(1-p)^{x-4}=\sum_{x\in A}\frac{(x-1)!}{3!(x-4)!}\left(1/3\right)^4\left(2/3\right)^{x-4},$$

$$\mathrm{E}(X)=\frac{4}{1/3}=12.$$

29. a) The memoryless property implies that the exponential density is appropriate but it requires $\sigma_X^2=[\mathrm{E}(X)]^2$. But, the given values do not suggest an exponential distribution. If we assume that the statistics produced by engineers are correct we have to ignore the memoryless characteristic and model using a gamma density. An appropriate probability model is,

$$\{R(X),\Upsilon,P(A)\}, R(X)=(0,\infty), \Upsilon=\{A: A \text{ is a Borel subset in } R(X)\},\ P(A)=\int_{x\in A} 4xe^{-2x}I_{(0,\infty)}(x)$$

b) No. When $\alpha=2>1$ gamma density exhibits "wear-out" effects since $P(x>s+t\,|\,x>s)$ declines as $s$ increases for $t>0$.

c) $$P\left(x<\frac{24\times 365.25}{100{,}000}\right)=P(x<.08766)=\int_0^{.08766}4xe^{-2x}dx=1-1.17532e^{-.17532}=.013685.$$

d) $P\left(x \geq \frac{4 \times 365.25 \times 5}{100,000}\right) = P\left(x \geq .07305\right) = \int_{0.07305}^{\infty} 4xe^{-2x}dx = 1.1461e^{-.1461} = .990312$.

31. a) $A: 82.255 - 100$
$B: 77.697 - 82.255$
$C: 72.303 - 77.697$
$D: 67.745 - 72.303$
$F: 0 - 72.303$

b) Normal distribution can only be an approximation since the grade distribution is contained within (0,100) while the normal distribution has a support spanning the entire real line.

33. a) $3.50 \pm 1.96\sqrt{.01} = \left[3.304, 3.696\right]$.

b) $P(Q > 110) = P\left(Z > \frac{110-100}{\sqrt{100}}\right) = P(Z > 1) = .1587$.

c) $M_X(\mathbf{t}) = \exp\left[\boldsymbol{\mu}'\mathbf{t} + \left(\frac{1}{2}\right)\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right], \mathbf{t} = \begin{bmatrix} t_P \\ t_Q \end{bmatrix} = \exp\left[3.5t_P + 100t_Q + .005t_P^2 - .7t_Pt_Q + 50t_Q^2\right]$,

$$\mathrm{E}(PQ) = \left.\frac{\partial M_X(\mathbf{t})}{\partial t_P \partial t_Q}\right|_{\mathbf{t}=\mathbf{0}} = \exp\left[\boldsymbol{\mu}'\mathbf{t} + \left(\frac{1}{2}\right)\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right], \mathbf{t} = \begin{bmatrix} t_P \\ t_Q \end{bmatrix} = 3.5 \times 100 - .7 = 349.3 = \$349,300.$$

d) $\mathrm{E}(Q \mid P) = \mu_2 + \frac{\sigma_{21}}{\sigma_1^2}\left(P - \mu_1\right) = 100 - \frac{.7}{.01}\left(P - 3.5\right) = 345 - 70P$. The negatively slowing demand curve makes economic sense since organic milk is a normal good.
$\mathrm{E}(Q \mid P = 3.40) = 345 - 70 \times 3.4 = 107{,}000$ gallons.

e) Each of the two variables, price and quantity can only take a non-negative value in reality. But, bivariate normally distributed random variables can take any real value. Therefore, a multivariate normal distribution is only an approximation in this case.

**Chapter 5 – Student Answer Key – Odd Numbers Only**
**Basic Asymptotics**

1. a) The marginal densities are

$$f_1(x_1;\alpha)=\int_0^1 \left(\alpha x_1+(2-\alpha)x_2\right)dx_2 I_{[0,1]}(x_1)=\left(\alpha x_1+\frac{1}{2}(2-\alpha)\right)I_{[0,1]}(x_1),$$

and

$$f_2(x_2;\alpha)=\int_0^1 \left(\alpha x_1+(2-\alpha)x_2\right)dx_1 I_{[0,1]}(x_2)=\left(\frac{\alpha}{2}+(2-\alpha)x_2\right)I_{[0,1]}(x_2),$$

with expectations

$$\mathrm{E}(X_1)=\int_0^1 x_1\left(\alpha x_1+\frac{1}{2}(2-\alpha)\right)dx_1=\frac{\alpha}{3}+\frac{(2-\alpha)}{4} \;,$$

and

$$\mathrm{E}(X_2)=\int_0^1 x_2\left(\frac{\alpha}{2}+(2-\alpha)x_2\right)dx_2=\frac{\alpha}{4}+\frac{2-\alpha}{3} \;.$$

For $\alpha\in(0,2)$ both $\mathrm{E}(X_1)<\infty$ and $\mathrm{E}(X_2)<\infty \Rightarrow \bar{X}_n \xrightarrow{\text{as}} \mu\left(\text{where } \bar{X}_n=\begin{bmatrix}\bar{X}_{1n}\\ \bar{X}_{2n}\end{bmatrix}\right)$ by Kolmogorov's SLLN (Theorem 5.26). Hence,

$$\bar{X}_n \xrightarrow{\text{as}} \mu \Rightarrow \bar{X}_n \xrightarrow{\text{p}} \mu \Rightarrow \bar{X}_n \xrightarrow{\text{d}} \mu.$$

b) From the multivariate LLCLT (Theorem 5.37)

$$Z_n=n^{1/2}\left[\bar{X}_n-\mu\right]\xrightarrow{\text{d}} N\left([0],\Sigma\right),$$

where

$$\mu=\begin{bmatrix}\frac{\alpha}{3}+\frac{(2-\alpha)}{4}\\ \frac{\alpha}{4}+\frac{(2-\alpha)}{3}\end{bmatrix} \text{ and } \Sigma\begin{bmatrix}\mathrm{var}(X_1) & \mathrm{cov}(X_1,X_2)\\ \mathrm{cov}(X_2,X_1) & \mathrm{var}(X_2)\end{bmatrix},$$

with

$$\operatorname{var}\left(X_1\right)=\left(\frac{\alpha}{4}+\frac{(2-\alpha)}{6}\right)-\left(\frac{\alpha}{3}+\frac{(2-\alpha)}{4}\right)^2,$$

$$\operatorname{var}\left(X_2\right)=\left(\frac{\alpha}{6}+\frac{(2-\alpha)}{4}\right)-\left(\frac{\alpha}{4}+\frac{(2-\alpha)}{3}\right)^2,$$

$$\operatorname{cov}\left(X_1,X_2\right)=\frac{1}{3}-\left(\frac{\alpha}{3}+\frac{2-\alpha}{4}\right)\left(\frac{\alpha}{4}+\frac{2-\alpha}{3}\right),$$

$$\Rightarrow \bar{X}_n=\frac{Z_n}{n^{1/2}}+\mu \overset{a}{\sim} N\left(\mu,n^{-1}\Sigma\right).$$

If $\alpha = 1$, $n = 200$, $\Sigma=\begin{bmatrix} .0764 & -.0069 \\ -.0069 & .0764 \end{bmatrix}$ and applying Theorem 4.11

$$P\left(\bar{X}_{1n}>.70 \mid \bar{X}_{2n}=.60\right)\approx\int_{.70}^{1.0} N\left(Z;.5818,.0758\right)dz=\int_{.4293}^{1.5190} N\left(Z;0,1\right)dz=.2693\,.$$

c) aslim $g\left(\bar{X}_n\right)=$ (aslim $\bar{X}_{1n}$)/(aslim $\bar{X}_{2n}$) $= \mu_1/\mu_2$ (Theorem 5.17)

$$\Rightarrow g\left(\bar{X}_n\right)\xrightarrow{p}\mu_1/\mu_2 \Rightarrow g\left(\bar{X}_n\right)\xrightarrow{d}\mu_1/\mu_2\,.$$

d) From Theorem 5.39,

$$\left.\frac{\partial g\left(\bar{x}_n\right)}{\partial \bar{x}_{1n}}\right|_{\bar{x}_n=\mu}=\left.\frac{1}{\bar{x}_{2n}}\right|_{\bar{x}_{2n}=.5833}=1.7143\neq 0,$$

$$\left.\frac{\partial g\left(\bar{x}_n\right)}{\partial \bar{x}_{2n}}\right|_{\bar{x}_n=\mu}=-\left.\frac{\bar{x}_{1n}}{\left(\bar{x}_{2n}\right)^2}\right|_{\substack{\bar{x}_n=.5833\\ \bar{x}_{2n}=.5833}}=-1.7143\neq 0,$$

$$\Rightarrow n^{1/2}\left(g\left(\bar{X}_n\right)-g\left(\mu\right)\right)\xrightarrow{d}N\left(0,G\Sigma G'\right)\Rightarrow g\left(\bar{X}_n\right)\overset{a}{\sim}N\left(g\left(\mu\right),n^{-1}G\Sigma G'\right),$$

where $g(\mu) = 1$, $\Sigma$ is defined in part (b), and $G = [1.7143, -1.7143]$.

Then, $P(g(\bar{x}_n) > 1)$ . $\int_1^\infty N(z;\,1,\,.0024)\,dz=\frac{1}{2}$

3. a) Both statements are true. First, considering convergence in mean square, since $\theta$ is a constant we note the following necessary and sufficient conditions for $Y_n \xrightarrow{m} \theta$ (Corollary 5.2):

i) $\mathrm{E}Y_n$ 6 $\theta$, and ii) $\operatorname{var}(Y_n)\to 0$.

Notice, $\mathrm{E}Y_n = n^{-1}\left(1+\sum_{i=1}^{n}\mathrm{E}X_i\right) = n^{-1}+\theta \to \theta$, meeting condition (i). Further,

$\operatorname{var}(Y_n) = \operatorname{var}\left(n^{-1}\left(1+\sum_{i=1}^{n}X_i\right)\right) = n^{-2} \operatorname{var}\left(\sum_{i=1}^{n}X_i\right) = n^{-1}\theta^2 \to 0$, meeting condition (ii).

Therefore, $Y_n \xrightarrow{m} \theta$. From Theorem 5.13 we know $Y_n \xrightarrow{m} \theta \Rightarrow Y_n \xrightarrow{p} \theta$.

b) To define an asymptotic distribution for $Y_n$, notice that $Y_n$ may be written as $Y_n = n^{-1} + \bar{X}_n$. From the Lindberg-Levy CLT (Theorem 5.30) we know that

$$\frac{n^{1/2}\left(\bar{X}_n-\theta\right)}{\theta} \xrightarrow{d} N(0,1),$$

so that

$$\bar{X}_n \overset{a}{\sim} N\left(\theta, \theta^2/n\right) \text{ and } Y_n \overset{a}{\sim} N\left(\theta+n^{-1}, \theta^2/n\right).$$

c) Under the stated conditions,

$Y_n \overset{a}{\sim} N(10.01,\ 1)$. Then $P(y_n \geq 15)$ . $P(z \geq 4.99) \approx 0$, where $Z \sim N(0,1)$.

5. a) **<u>NOTE</u>:** The definition of $\bar{X}_d$ should be $\bar{X}_d = (1/d)\sum_{t=1}^{d} X_t/N_t$. To show $\bar{X}_d \xrightarrow{as} p$, we first show that the $X_t^* = X_t/N_t$ are asymptotically nonpositively correlated, and then use the SLLN (Theorems 5.28 and 5.29) to prove $\bar{X}_d \xrightarrow{as} p$.

By Theorems 3.11 and 3.12

$$\mathrm{E}\left(X_t^* \mid n_t\right) = \mathrm{E}\left(\frac{X_t}{n_t}\middle| n_t\right) = p \Rightarrow \mathrm{E}\left(X_t^*\right) = \mathrm{E}\left(\mathrm{E}\left(X_t^* \mid N_t\right)\right) = p.$$

Similarly,

$$\mathrm{E}\left(X_t^{*2} \mid n_t\right) = \left(\frac{X_t^2}{n_t^2}\middle| n_t\right) = \frac{1}{n_t^2}\left(n_t p(1-p) + (n_t p)^2\right) = \frac{p(1-p)}{n_t} + p^2,$$

$$\Rightarrow \mathrm{E}\left(X_t^{*2}\right) = \mathrm{E}\left(\mathrm{E}\left(X_t^{*2} \mid N_t\right)\right) = \left(p-p^2\right)\mathrm{E}\left(\frac{1}{N_t}\right) + p^2 = \left(p-p^2\right)(.004056) + p^2.$$

**<u>NOTE</u>**: $\mathrm{E}\left(\frac{1}{N_t}\right) = \sum_{n_t=200}^{300}\left(\frac{1}{n_t}\right)\left(\frac{1}{101}\right)$.

Since $\operatorname{var}\left(X_t^*\right) = \mathrm{E}\left(X_t^{*2}\right) - \left(\mathrm{E}X_t^*\right)^2$, it is bounded.

Because the $X_t$'s and $N_t$'s are all independent, then the $X_t^*$'s are independent, which yields $\text{cov}\left(X_i^*, X_{i+t}^*\right) = 0 \forall t$. Consequently, for any convergent series, $\sum_{i=`}^{\infty} a_i < \infty$, $\text{cov}\left(X_t^*, X_{i+t}^*\right) \leq a_t \sigma_i \sigma_{i+t}$. Hence, the sequence of random scalars $\left\{X_t^*\right\}$ are asymptotic nonpositively correlated.

By Theorems 5.28 and 5.29, $\bar{X}_d \xrightarrow{as} p$, which implies $\bar{X}_d \xrightarrow{p} p$ and $\bar{X}_d \xrightarrow{d} p$.

b) By LLCLT
$\bar{X}_d \overset{a}{\sim} N\left(\mu, \sigma^2 / d\right)$.

Thus, with $\mu$=$p$ and $\sigma^2$ = ($p$−$p^2$)(.004056),

$$P\left(.78 < \bar{x}_d < .82\right) = P\left(\frac{.78 - p}{\sqrt{\left(p - p^2\right)\left(\frac{.004056}{d}\right)}} < z < \frac{.82 - p}{\sqrt{\left(p - p^2\right)\left(\frac{.004056}{d}\right)}}\right),$$
$$= P\left(-13.61 < z < 13.61\right) \approx 1.$$

7. a) $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$, with $M_{X_i}\left(t\right) = \left(1 - 3t\right)^{-2} \forall i$. From Theorem 3.27, p.144,

$$M_{\bar{X}}\left(t\right) = \prod_{i=1}^{n} M_{X_i}\left(t / n\right) = \left(1 - \frac{3t}{n}\right)^{-2n} \Rightarrow \bar{X} \sim \text{Gamma}\left(2n, 3 / n\right).$$

b) By the LLCLT, $\bar{X}_n \overset{a}{\sim} N\left(\alpha\beta, \alpha\beta^2 / n\right) = N\left(6, 18 / n\right)$.

c) This will be a plot of Gamma(20, .3) and $N$(6, 1.8).

d) This will be a plot of Gamma (80, .075) and $N$(6, .45).

The graphs when $n$=40 will look much more similar than when $n$=10. This is reflecting the improved approximation accuracy of the asymptotic normal density as n increases.

9. Assume the responses are outcomes of iid Bernoulli trials. By the LLCLT

$$\bar{X}_n \overset{a}{\sim} N\left(p, \frac{\left(1 - p\right)p}{n}\right),$$

so,

$$P\left(\left|\bar{x}_n-\mu\right|<.02\right)=P\left(\left|\frac{\bar{x}_n-p}{\sqrt{\frac{(1-p)p}{n}}}\right|\le\frac{.02}{\sqrt{\frac{(1-p)p}{n}}}\right)=.99$$ , and since $P(-2.575 \le z \le 2.575)=.99$

$$\Rightarrow n=\left(\frac{2.575}{.02}\right)^2\ (p(1-p)).$$

Since $p(1-p)$ is maximized at .5, choose $p$ = .5 to make sure the probability statements holds $\Rightarrow n=4144$.

11. a) Using the LLCLT, and the fact that $\mathrm{E}X_i$ = .5 and $\mathrm{var}(X_i)$ = 1/12, $\bar{X}_n \overset{a}{\sim} N\left(.5,\frac{1}{12n}\right)$.

b) From a), $\sum_{i=1}^n X_i=n\bar{X}_n \overset{a}{\sim} N\left(.5n,\frac{n}{12}\right)$. Then for $n$=12, $\sum_{i=1}^{12} X_i \overset{a}{\sim} N(6,1)$, and so $\sum_{i=1}^{12} X_i-6 \overset{a}{\sim} N(0,1)$.

13. a) $\bar{x}=n^{-1}\sum_{i=1}^n x_i=\frac{29{,}200}{1{,}460}=20$ tons.

b) By the LLCLT

$$Z_n=\sqrt{n}\frac{\left(\bar{X}_n-\mu\right)}{\sigma}\xrightarrow{d}N(0,1)\Rightarrow\bar{X}_n=\frac{\sigma}{\sqrt{n}}Z_n+\mu \overset{a}{\sim} N\left(\mu,\frac{\sigma^2}{n}\right).$$

For the triangular distribution, $\mu=a$ and

$$\mathrm{var}(X)=\mathrm{E}\left(X^2\right)-a^2=a^2+2/3-a^2=2/3,$$

where

$$\mathrm{E}\left(X^2\right)=\int_{a-2}^{a}x^2\left(\left(\frac{1}{2}-\frac{1}{4}a\right)+\frac{1}{4}x\right)dx+\int_{a}^{a+2}x^2\left(\left(\frac{1}{2}+\frac{1}{4}a\right)-\frac{1}{4}x\right)dx=a^2+2/3.$$

Hence, $\bar{X}_n \overset{a}{\sim} N\left(a,\frac{2}{3(1460)}\right)=N(a,\ .000457)$.

c) $P\left(|x_n - a| < .05\right) = P\left(\frac{|x_n - a|}{\sqrt{.000457}} < \frac{.05}{\sqrt{.000457}}\right) \approx \int_{-2.34}^{2.34} N(z;0,1)\,dz = .981.$

d) Using the Berry-Esséen inequality (see pages 264-265 and footnote 11)

$$\max_z |F_n(z) - F(z)| \le \frac{.7975\ \mathrm{E}\left[|X-\mu|^3\right]}{n^{1/2}\sigma^3} = \frac{7975(.8)}{(1460)^{1/2}(2/3)^{3/2}} = .0307.$$

Yes, with a conservative approximation error of .0307 and with probability .981 that $|\bar{x} - \mu| < .05$, we have a reasonably accurate estimate. Note, the probability of $|\bar{x} - \mu| < 1$ would be $\approx 1$ based on the asymptotic distribution.

Note that

$$\mathrm{E}\left[|x-a|^3\right] = \int_{a-2}^{a} (a-x)^3\left(\left(\frac{1}{2} - \frac{1}{4}a\right) + \frac{1}{4}x\right)dx + \int_{a}^{a+2} (x-a)^3\left(\left(\frac{1}{2} + \frac{1}{4}a\right) - \frac{1}{4}x\right)dx.$$

Substituting $\mu = a - x$ in the first integral and $\mu = x - a$ in the second integral obtains

$$\mathrm{E}\left(|X-a|^3\right) = -\int_{2}^{0} \mu^3\left(\left(1/2 - \frac{\alpha}{4}\right) + \frac{a-\mu}{4}\right)d\mu + \int_{0}^{2} \mu^3\left(\left(1/2 + \frac{a}{4}\right) - \frac{\mu+a}{4}\right)d\mu$$
$$= \int_{0}^{2} \mu^3\left[1 - \frac{\mu}{2}\right]du = \left.\frac{\mu^4}{4} - \frac{\mu^5}{10}\right|_0^2 = .8.$$

e) $f(x;\hat{a}) = [-4.5 + .25x]\,I_{[18,20]}(x) + [5.5 - .25x]\,I_{(20,22]}(x),$

$$P(x > 21) = \int_{21}^{22} (5.5 - .25x)\,dx = .1250.$$

15. From the std normal limiting distribution result, we can deduce that $Y_n \overset{a}{\sim} N(n, 2n)$. Then

| | | Actual | Asymptotic $N(n, 2n)$ |
|---|---|---|---|
| $P(y_{25} \le 34.3816)$ | = | .90 | .9077 |
| $P(y_{50} \le 63.1671)$ | = | .90 | .9060 |
| $P(y_{100} \le 118.498)$ | = | .90 | .9046 |

These are all within .008 of the true value--reasonably good.

$$\mathrm{E}(Y_n/n) = 1,\ \mathrm{var}(Y_n/n) = 2/n \to 0 \Rightarrow Y_n/n \xrightarrow{m} 1 \Rightarrow Y_n/n \xrightarrow{p} 1.$$

17. a) Since $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$, we have the estimate, $\hat{p} = 49{,}873/100{,}000 = .49873$.

b) Since the spin outcomes are viewed as iid Bernoulli, we have from the Lindberg-Levy CLT that
$\dfrac{n^{1/2}\left(\bar{X}_n - p\right)}{\left(p(1-p)\right)^{1/2}} \xrightarrow{d} N(0,1)$, from which it follows that $\bar{X}_n \overset{a}{\sim} N\left(p,\, p(1-p)/n\right)$.

c) Using $\hat{p} = .49873$, we have that $\bar{X}_n \overset{a}{\sim} N(.49873,\, .0000025)$.

Then

$$P\left[\left|\bar{x}_n - .49873\right| \le .005\right] = P\left[-.005 \le \bar{x}_n - .49873 \le .005\right]$$
$$= P\left[-.005/(.0000025)^{1/2} \le z \le .005/(.0000025)^{1/2}\right] = F(3.16) - F(-3.16) = .9984.$$

where $Z \sim N(0,1)$ and $F(\cdot)$ is the standard normal CDF.

From Chebyshev's inequality we have

$$P\left[\left|\bar{x}_n - p\right| < k\left(\frac{p(1-p)}{100{,}000}\right)^{1/2}\right] \ge 1 - \frac{1}{k^2}.$$

For any given $k$, the upper bound will be its largest when $p$=.5. So considering this worst case scenario (in terms of the variance of $\bar{X}_n$ being the largest), rewrite the inequality with $p$=.5 and, for example, $k$=5, to obtain

$$P\left[\left|\bar{x}_n - .5\right| < .0079\right] \ge .96.$$

So given $p$=.5, we see that the probability that an outcome of $\bar{X}_n$ is within $\pm$.0079 of the true probability ($p$) is greater than or equal to .96.

d) Given the $X_i$'s are iid Bernoulli, then, under conditions of the Lindberg-Levy CLT, it can be shown that the maximum absolute difference between $P(y_n \le c)$ for the actual density of $Y_n = n^{1/2}\left(\bar{X}_n - \mu\right)/\sigma$ and for that of the standard normal is

$$.7975n^{-1/2}\left[1 - 2p(1-p)\right]\left[p(1-p)\right]^{-1/2}.$$

Using the outcome of $\bar{X}_n$ (i.e., .49873) as an estimate of $p$, the bound is equal to .0025 when $n$=100,000.

e) By the Lindberg-Levy CLT,

$$Z_n = \frac{n^{1/2}\left(\bar{X}_n - p\right)}{\left(p(1-p)\right)^{1/2}} \xrightarrow{d} Z \sim N(0,1).$$

Define $y_n = \left(p(1-p)\right)^{1/2}\ \forall n$ which is some constant depending upon the value $p$. We may then rely upon Slutsky's Theorems (Theorem 5.10) to define a new sequence and limiting distribution as

$$y_n Z_n \xrightarrow{d} cZ \text{ where } c = p(1-p)^{1/2}, \text{ or } n^{1/2}\left(\bar{X}_n - p\right) \xrightarrow{d} N\left(0,\ p(1-p)\right).$$

Theorem 5.39 now applies with $\mu = p$ and $\Sigma = \sigma^2 = p(1-p)$. Define $g\left(\bar{X}_n\right) = \bar{X}_n\left(1-\bar{X}_n\right)$, and notice that $dg(p)/d\bar{x} = 1-2p$, which is continuous in $p$ and is nonzero as long as $p \neq .5, 0, \text{or } 1$. Then for $p \neq 0, .5, \text{ or } 1$, Theorem 5.39 implies that

$$n^{1/2}\left(\bar{X}_n\left(1-\bar{X}_n\right) - p(1-p)\right) \xrightarrow{d} N\left(0, (1-2p)^2\ p(1-p)\right),$$

and

$$\bar{X}_n\left(1-\bar{X}_n\right) \overset{a}{\sim} N\left(p(1-p),\ n^{-1}(1-2p)^2\ p(1-p)\right).$$

For the case where $p$=.5, see Example 5.46 for details. For $p$=0 or $p$=1, the random variable is degenerate.

f) First of all, one can derive the limiting distribution of $n^{1/2}\left(S_*^2 - \sigma^2\right)$, where $S_*^2 = n^{-1}\sum_{i=1}^n\left(X_i - \bar{X}_n\right)^2$, by following an argument similar to the one on p. 317-318, Chapter 6, leading to

$$Z_n = n^{1/2}\left(S_*^2 - \sigma^2\right) \xrightarrow{d} Z \sim N\left(0, \mu_4 - \sigma^4\right).$$

The statement remains true if $S^2$ replaces $S_*^2$, since by Slutsky's theorem

$$n^{1/2}\left(S^2 - \sigma^2\right) = \left(\frac{n}{n-1}\right)Z_n + \frac{n^{1/2}}{n-1}\sigma^2 \xrightarrow{d} Z \sim N\left(0, \mu_4 - \sigma^4\right).$$

Where $\left(\frac{n}{n-1}\right) \to 1$ and $\frac{n^{1/2}}{n-1}\sigma^2 \to 0$. Then

$$S^2 \overset{a}{\sim} N\left(p(1-p), n^{-1}p(1-p)(1-2p)^2\right),$$

where $\mu_4 = \mathrm{E}\left(X_i - p\right)^4 = (1-p)^4 p + (0-p)^4 (1-p) = p(1-p)\left(3p^2 - 3p + 1\right),$

and

$$\sigma^4 = \left[p(1-p)\right]^2 = p^2 - 2p^3 + p^4.$$

Since the asymptotic distributions of both estimators are the same, they can=t be distinguished on this basis.

***INSTRUCTOR NOTE**: It is possible to derive other asymptotic distributions for $S^2$ since asymptotic distributions are not unique. See discussion on p. 388 relating to CAN class of estimators and a justification for the choice of centering used in deriving the asymptotic distribution of $S^2$ above.

19. First note that

$$\text{plim}\left(S_n^2\right) = \text{plim}\left(n^{-1}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2\right) = \lambda(1+\lambda) - \lambda^2 = \lambda$$

since the $X_i^2$'s are iid with $\mathrm{E}X_i^2 = \lambda(1+\lambda)\,\forall i$ and Khinchin's WLLN applies to the first summation term and Khinchin's WLLN also applies as $\bar{X}_n \xrightarrow{p} \lambda$ and thus $\bar{X}_n^2 \xrightarrow{p} \lambda^2$ by Theorem 5.5.

From the LLCLT, and since $\mu$=$\lambda$ and $\sigma^2$=$\lambda$,

$$Z_n = \frac{\bar{X}_n - \lambda}{(\lambda/n)^{1/2}} = \frac{n^{1/2}\left(\bar{X}_n - \lambda\right)}{\lambda^{1/2}} \xrightarrow{d} N(0,1).$$

From Theorem 5.5 and the previous result, $\left(\lambda / S_n^2\right)^{1/2} \xrightarrow{p} 1,$ so that by Slutsky's theorem

$$\left(\lambda / S_n^2\right)^{1/2} Z_n = \frac{n^{1/2}\left(\bar{X}_n - \lambda\right)}{\left(S_n^2\right)^{1/2}} = \frac{\bar{X}_n - \lambda}{\left(S_n^2/n\right)^{1/2}} \xrightarrow{d} N(0,1).$$

**Chapter 6 – Student Answer Key – Odd Numbers Only**
**Sampling, Sample Moments, Sampling**
**Distributions, and Simulation**

1. a) Dividing the numerator and denominator of $T$ by $\sigma$, we have

$$T = \frac{n^{1/2}\left(\bar{X}-\mu\right)/\sigma}{\hat{\sigma}/\sigma}.$$

The numerator of $T$ is a $N(0,1)$ random variable. Using the definition of $\hat{\sigma}$ given in Theorem 6.19, we have

$$\hat{\sigma}/\sigma = \left(\left(nS^2/\sigma^2\right)/\left(n-1\right)\right)^{1/2},$$

i.e., the denominator of $T$ is the square root of a chi-square random variable ($nS^2/\sigma^2$) divided by its degrees of freedom ($n-1$). Finally, since by Theorem 6.12 $\bar{X}$ and $S^2$ are independent random variables we have, by Theorem 6.18, that $T$ has the $t$-density with $n-1$ degrees of freedom.

b)

$$\begin{aligned} P\left(\bar{x}-2.06\hat{\sigma}/n^{1/2} < \mu < \bar{x}+2.06\hat{\sigma}/n^{1/2}\right) &= P\left(-2.06 < n^{1/2}\left(\mu-\bar{x}\right)/\hat{\sigma} < 2.06\right) \\ &= P\left(-2.06 < n^{1/2}\left(\bar{x}-\mu\right)/\hat{\sigma} < 2.06\right) \\ &= P\left(-2.06 < t < 2.06\right) = .95. \end{aligned}$$

where $T$ has the $t$-density with 24 degrees of freedom.

c) With $n$=25, the 90% confidence interval is $\left(\bar{X}-1.711\hat{\sigma}/5,\ \bar{X}+1.711\hat{\sigma}/5\right)$.

Since $\hat{\sigma} = \left(nS^2/\left(n-1\right)\right)^{1/2} = .1021$ the confidence interval outcome is (16.265, 16.335).

3. a) $\dfrac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$ (Theorem 6.12), so that

$$P\left(\chi^2_{1-\alpha} < \frac{ns^2}{\sigma^2} < \chi^2_{\alpha}\right) = P\left(\frac{ns^2}{\chi^2_{\alpha}} < \sigma^2 < \frac{ns^2}{\chi^2_{1-\alpha}}\right) = 1-2\alpha.$$

b) For 19 degrees of freedom, $\chi^2_{.025}$ = 32.8523 and $\chi^2_{.975}$ = 8.90655. Then the confidence interval is

$\left(\frac{20s^2}{32.8523} < \sigma^2 < \frac{20s^2}{8.90655}\right)$ with outcome (5.479, 20.210).

5. a) Composite experiment. Note the expected values of the $Y_j$'s will generally differ.

b)

$$\mathrm{E}\bar{Y}_n = \beta_0 + \beta_1 \sum_j f_j / n + \beta_2 \sum_j f_j^2 / n, \quad n = 40,$$

$$\operatorname{var} \bar{Y}_n = \sigma^2 / n, \quad n = 40,$$

$$\mathrm{E}\left[\bar{Y}_n - \mathrm{E}\bar{Y}_n\right] = 0 \text{ and } \operatorname{var}\left(\bar{Y}_n\right) \to 0 \text{ as } n \to \infty.$$

$$\Rightarrow \bar{Y}_n - \mathrm{E}\bar{Y}_n \xrightarrow{p} 0.$$

Then $\bar{Y}_n$ will converge in probability to a constant iff $\mathrm{E}\bar{Y}_n \to c$.

c) Yes, by the results in part b). The probability that an outcome of $\bar{Y}_n$ deviates from $\mathrm{E}\bar{Y}_n$ by any nonzero amount goes to zero as $n \to \infty$.

The outcome of $\bar{Y}_n$ will be a meaningful estimate of the simple average of the expected yields across all one-acre plots.

d) Yes. The $V_j$'s represent influences on the $Y_j$'s other than fertilizer. Being contiguous, the plots could all be affected by common influences like pest infestation, temperature variation, rainfall, etc., negating the iid assumption. Won't change answer to part a)--still a composite experiment.

7. a) Notice that $q = g(v) = c(\mathbf{z})v$ is a monotonically increasing function of $v$ and thus is invertible where $v = g^{-1}(q) = q/c(\mathbf{z})$. Applying Theorem 6.15 we have

$$h(q) = \theta\left(\frac{q}{c(\mathbf{z})}\right)^{\theta-1}\left(\frac{1}{c(\mathbf{z})}\right) I_{(0,c(\mathbf{z}))}(q).$$

b)

$$\mathrm{E}q = \int_{-\infty}^{\infty} q h(q)\,dq = \int_{-\infty}^{\infty} \frac{10q}{5000}\left(\frac{q}{5000}\right)^9 I_{(0,5000)}(q)\,dq,$$

$$= \frac{10}{5000^{10}} \int_0^{5000} q^{10} dq = 4{,}545.45 \text{ tons of fish}$$

9. a) By Theorem 6.15, with $v = g(\varepsilon) = e^{\varepsilon}$,

$$h(v) = \begin{cases} f\left(g^{-1}(v)\right)\left|\dfrac{dg^{-1}(v)}{dv}\right|, & \text{for } v \in \Psi = \{v : v = e^{\varepsilon}, \varepsilon \in \Xi\} \text{ and } \Xi = \{\varepsilon : f(\varepsilon) > 0\} = (-\infty, \infty) \\ 0 & \text{, otherwise} \end{cases}$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{\ln(v)}{\sigma}\right)^2\right)\cdot\left|\frac{1}{v}\right| I_{(0,\infty)}(v)$$

$$= \frac{1}{\sqrt{2\pi}\sigma v}\exp\left(-\frac{1}{2}\left(\frac{\ln(v)}{\sigma}\right)^2\right) I_{(0,\infty)}(v).$$

b) If $x_1 = x_2 = 2$, then $Q = 18.025\ V$. So, with $q = g(v) = 18.025v$,

$$d(q) = \begin{cases} f\left(g^{-1}(q)\right)\left|\dfrac{dg^{-1}(q)}{dq}\right| & \text{, for } q \in \Psi = \{q : q = 18.025v,\ v \in (0,\infty)\} \\ 0 & \text{, otherwise} \end{cases}$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{q/18.025}\exp\left(-\frac{1}{2}\left(\frac{\ln\left(\frac{q}{18.025}\right)}{\sigma}\right)^2\right)\left|\frac{1}{18.025}\right| I_{(0,\infty)}(q)$$

$$= \frac{1}{\sqrt{2\pi}\sigma q}\exp\left(-\frac{1}{2}\left(\frac{\ln(q) - 2.8918}{\sigma}\right)^2\right) I_{(0,\infty)}(q).$$

c) In general,

$$\mathrm{E}Q = 10x_1^{.35}x_2^{.5}\mathrm{E}V = 10x_1^{.35}x_2^{.5}\exp\left(\sigma^2/2\right) = 29.72 \ \left(\text{when } x_1 = x_2 = 2 \text{ and } \sigma^2 = 1\right).$$

d) For $i$=1,...,1600, $\mathrm{E}V_i = e^{1/8} = 1.1331$ and $\mathrm{var}(V_i) = e^{1/2} - e^{1/4} = .3647$.

Since the $V_i$'s are iid r.v.'s, then by LLCLT

$$Z_n = \frac{\sum_{i=1}^n V_i - n(1.1331)}{n^{1/2}\sqrt{.3647}} \xrightarrow{d} N(0,1).$$

$$\Rightarrow Q_n^* = \sum_{i=1}^n Q_i = 10x_1^{.35}x_2^{.5}\left[n^{1/2}\sqrt{.3647}Z_n + (1.1331)n\right].$$

Letting $x_1 = 7.24579$, $x_2 = 4$, $n = 1600 \Rightarrow Q^* \overset{a}{\sim} N(72{,}518.40,\ 933{,}619.74)$. Yes, $Q^*$ can be considered approximately normally distributed.

e) The upper bound on the approximation error is

$$\frac{.7975\ \mathrm{E}|V_i - \mathrm{E}V_i|^3}{\sigma^3 n^{1/2}} = \frac{(.7975)(.4010)}{(1600)^{1/2}(.6039)^3} = .0363.$$

11. a) From Problem 1, part a), we know that $T = \left(\bar{X} - \mathrm{E}\bar{X}\right) / \left(\hat{\sigma}_X^2 / n\right)^{1/2}$ has a *t*-distribution with ($n-1$) degrees of freedom. Then $P\left(|t| \leq 1.316\right) = .80$ (Table B.2, 25 *df*).

b) From Theorem 6.12, $26S_X^2 / \sigma_X^2$ as a $\chi^2$ distribution with 25 *df*. Then $P\left(26s_X^2 / \sigma_X^2\right) > 37.652 = .05$ (Table B.3).

c) $P\left(s_X^2 > 6.02432\right) = P\left(\frac{26s_X^2}{4} \geq 39.158\right) = .0355$ (linearly interpolating in Table B.3).

d)

$$P\left(s_Y^2 > 1.92s_X^2\right) = P\left(s_Y^2 / s_X^2 > 1.92\right)$$

$$P\left(\frac{31s_Y^2 / 30}{26s_X^2 / 25} > 1.91\right) = P\left(F_{30,25} > 1.91\right) \approx .05 \text{ (from Table B.4).}$$

e)

$$P\left(\frac{\hat{\sigma}_Y^2 \sigma_X^2}{\hat{\sigma}_X^2 \sigma_Y^2} \geq c\right) = P\left(F_{30,25} \geq 1.9192\right) = .05$$

$$\Rightarrow P\left(\frac{\sigma_Y^2 \hat{\sigma}_X^2}{\sigma_X^2 \hat{\sigma}_Y^2} \leq \frac{1}{c}\right) = P\left(F_{25,30} \leq .5211\right) = .05.$$

13. a) Following Example 6.21, define

$$g(y) = \begin{cases} 0 \text{ if } 0 < y \leq .85, \\ 1 \text{ if } .85 < y \leq 1, \end{cases}$$

based on a Bernoulli density with $p = .15$, where 1 denotes a purchase, 0 denotes no purchase. The simulated random sample of consumer purchasing decisions is then $\{0,0,0,0,0,0,0,0,0,0\}$.

b) $\bar{x} = 0$ and $s^2 = 0$, whereas $\mu = .15$ and $\sigma^2 = .1275$.

15. a) For $\lambda = .1$ and $f(x) = \dfrac{e^{-.1}(.1)^x}{x!}$ for $x \epsilon \{0,1,2,...\}$.

| x | f(x) |
|---|---|
| 0 | .9048 |
| 1 | .0905 |
| 2 | .0045 |

Define $g(y)$ as

$$g(y) = \begin{cases} 0 & \text{if} \quad 0 < y \leq .9048, \\ 1 & \text{if} \quad .9048 < y \leq .9953, \\ 2 & \text{if} \quad .9953 < y \leq .9998. \end{cases}$$

then the simulated random sample of copy machine daily malfunctions is $\{0,0,0,0,0,1,0,0,0,0\}$.

b) The sample mean and variance are $\bar{x} = .1$ and $s^2 = .09$ while the population mean and variance are $\mu = .1$ and $\sigma^2 = .1$.

17. a) Using the hint (referring to Problem 6), let $X_1 \sim$ Gamma$(1,\beta)$ and $X_2 \sim$ Gamma$(20, \beta)$. Consider $X_1 \sim$ Gamma$(1, \beta)$ and note that (choosing $\beta$=1)

$$f(x_1;1,1) = e^{-x_1} I_{(0,\infty)}(x_1),$$

with CDF

$$F(x_1) = 1 - e^{-x_1}.$$

From Theorem 6.23, we have that $X_1 = F^{-1}(Y)$ or $X_1 = -\ln(1-Y)$ where $Y \sim I_{(0,1)}(y)$. Likewise, $X_2 \sim$ Gamma $(20,\beta)$ (and choosing $\beta$=1) yields

$$f(x_2;20,1) = \frac{1}{19!} x_2^{19} e^{-x_2} I_{(0,\infty)}(x_2),$$

with CDF

$$F(x_2) = \int_0^{x_2} \frac{1}{19!} z^{19} e^{-z} dz.$$

The most efficient method by which to solve for the $x_2$ values is using a computer.

| Uniform (0,1) | $x_1$ | Uniform (0,1) | $x_2$ | $y = x_1/(x_1+x_2)$ |
|---|---|---|---|---|
| .6829 | 1.1485 | .0801 | 14.0867 | .0754 |
| .4283 | .5591 | .6117 | 20.9511 | .0260 |
| .0505 | .0518 | .3706 | 18.2412 | .0028 |
| .7314 | 1.3145 | .2936 | 17.3606 | .0704 |
| .8538 | 1.9228 | .2799 | 17.1972 | .1006 |
| .6762 | 1.1276 | .3900 | 18.4567 | .0576 |
| .6895 | 1.1696 | .7533 | 22.8590 | .0487 |
| .9955 | 5.4037 | .0113 | 11.2213 | .3250 |
| .2201 | .2486 | .5659 | 20.4118 | .0120 |
| .9144 | 2.4581 | .9063 | 26.0982 | .0861 |
| .3982 | .5078 | .5029 | 19.6999 | .0251 |
| .9574 | 3.1559 | .6385 | 21.2795 | .1292 |

(**NOTE**: We used the Nonlinear Equations solver in GAUSS to generate the $x_2$ values.)

b) The population mean and variance with $\alpha$=1, $\beta$=20 are

$\mu = \alpha/(\alpha+\beta) = .0476$ and $\sigma^2 = \alpha\beta/\left[(\alpha+\beta+1)(\alpha+\beta)^2\right] = .0021.$

The sample mean and variance are

$\bar{y} = .0799$ and $s^2 = .0068.$

19. a) This is a step function with increments of 1/20 added at each of the sorted values of the 20 outcomes, and then extrapolated to $-\infty$ on the left and $+\infty$ on the right.

b) $\hat{P}(x > 26) =$ (number of outcomes > 26)/20 = .20.

c) $\hat{P}(x \in [24, 26]) =$ (number of outcomes ϵ[24, 26])/20 = .35.

d) This is equal to the sample mean, since the EDF assigns probability 1/20 to each of the 20 outcomes, and thus

$$\mathrm{E}_{\hat{F}_n}(X) = \frac{1}{20}\sum_{i=1}^{20} x_i = \bar{x} = 24.522.$$

21. a) $s_{YP} = 1/14\sum_{i=1}^{14}(y_i - \bar{y})(p_i - \bar{p}) = 1{,}666{,}513{,}647.02$.

b) $r_{YP} = \dfrac{s_{YP}}{s_Y s_P} = .9992$.

c) $\hat{p}_i = \overline{p} - \dfrac{s_{YP}}{s_Y^2}\overline{y} + \dfrac{s_{YP}}{s_Y^2} y_i = 19{,}075.68 + 1.54587 y_i$ .

d) $r_{YP}^2 = .9984$ . Thus 99.84% of the sample variance in home price is explained by the linear function of income in Part c).

23. a) $\sum_{i=1}^{3} X_i \sim \text{Gamma}\left(\sum_{i=1}^{3}\alpha_i, \beta\right)$, where $\sum_{i=1}^{3}\alpha_i = 3,\ \beta = 10.$ Hence,

$$P\left(\sum_{i=1}^{3} x_i \geq 20{,}000\right) = \int_{20}^{\infty} \frac{1}{(10)^3\,\Gamma(3)} z^2 e^{-z/10} dz = .6767$$

obtained using the computer, or else via integration by parts and L'Hospital's rule.

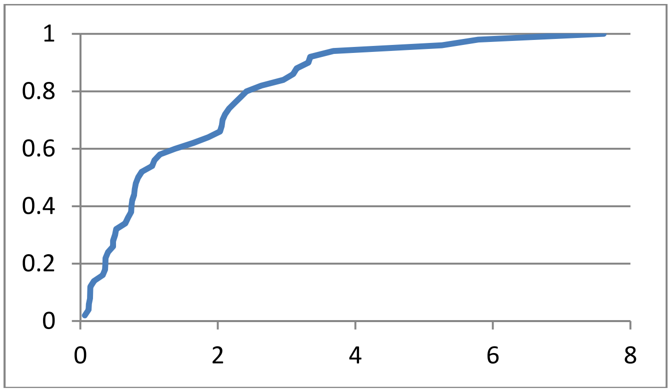
b) $P\left(\sum_{i=1}^{2} x_i \geq 20{,}000\right) = \int_{20}^{\infty} \frac{1}{(10)^2\,\Gamma(2)} z e^{-z/10} dz = .4060$ .

Reducing to two systems decreases the probability the navigation component will continue to function.

25. a)

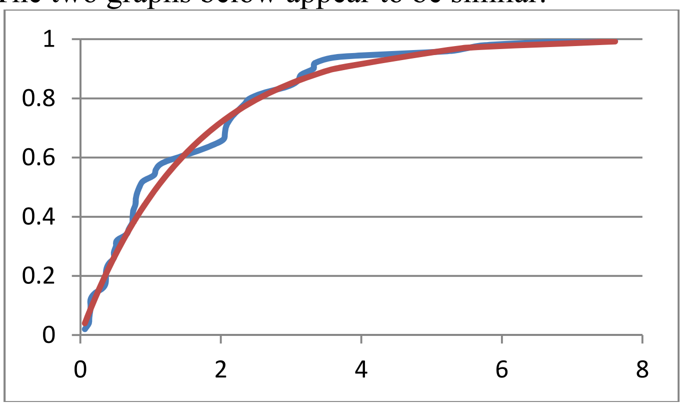
$$\begin{aligned}
F(x) = {} & .02I_{[.064,.121)}(x) + .04I_{[.121,126)}(x) + .06I_{[.126,.141)}(x) + .08I_{[.141,.145)}(x) + .10I_{(.145,.146)}(x) + .12I_{(.146,.200)}(x) \\
& + .14I_{[.200,.326)}(x) + .16I_{[.326,361)}(x) + .18I_{[.361,.366)}(x) + .20I_{[.366,.368)}(x) + .22I_{(.368,.402)}(x) + .24I_{(.402,.476)}(x) \\
& + .26I_{[.476,.478)}(x) + .28I_{[.478,.506)}(x) + .30I_{[.506,522)}(x) + .32I_{[.522,.654)}(x) + .34I_{[.654,.694)}(x) + .36I_{(.694,.740)}(x) \\
& + .38I_{[.740,.744)}(x) + .40I_{[.744,.753)}(x) + .42I_{[.753,.783)}(x) + .44I_{[.783,.790)}(x) + .46I_{(.790,.808)}(x) + .48I_{(.808,.841)}(x) \\
& + .50I_{[.841,.893)}(x) + .52I_{[.893,1.045)}(x) + .54I_{[1.045,1.077)}(x) + .56I_{[1.077,1.156)}(x) + .58I_{(1.156,1.381)}(x) + .60I_{(1.381,.1.641)}(x) \\
& + .62I_{[1.641,1.862)}(x) + .64I_{[1.862,2.031)}(x) + .66I_{[2.031,2.057)}(x) + .68I_{[2.057,2.068)}(x) + .70I_{(2.068,2.105)}(x) + .72I_{(2.105,2.165)}(x) \\
& + .74I_{[2.165,2.249)}(x) + .76I_{[2.249,2.334)}(x) + .78I_{[2.334,2.415)}(x) + .80I_{[2.415,2.631)}(x) + .82I_{(2.631,2.953)}(x) + .84I_{(3.953,3.096)}(x) \\
& + .86I_{[3.096,3.145)}(x) + .88I_{[3.145,3.317)}(x) + .90I_{[3.317,3.344)}(x) + .92I_{[3.344,3.675)}(x) + .94I_{(3.675,5.258)}(x) + .96I_{(5.258,.5.788)}(x) \\
& + .98I_{[5.788,7.611)}(x) + I_{[7.611,\infty)}(x)
\end{aligned}$$

b)




c) 1.57704.

d) The two graphs below appear to be similar.




e) The third moment about the mean implied by the EDF is 6.733132 which means that the PDF of screen lifetimes is not a symmetric probability distribution.

f) $P\left(x < \dfrac{365.25 \times 24}{100000}\right) = P(x < .08766) = .02$.

g) $P\left(x > \dfrac{5 \times 365.25 \times 4}{100000}\right) = P(x > .07305) = .98$.

27. a) $F(t) = .35 I_{[1,2)} + .75 I_{[2,3)} + .9 I_{[3,4)} + I_{[4,\infty)}$

b) $R(t) = \{1,2,3,4\},\ f(t) = .35 I_{\{1\}} + .4 I_{\{2\}} + .15 I_{\{3\}} + .1 I_{\{4\}}$.

c)

$$\mathrm{E}(t) = .35 + 2 \times .4 + 3 \times .15 + 4 \times .1 = 2,$$
$$\mathrm{var}(t) = \frac{20}{19}\left[(1-2)^2 \times .35 + (2-2)^2 \times .4 + (3-2)^2 \times .15 + (4-2)^2 \times .1\right] = .947368.$$

d)

$$\mathrm{E}(t) = \frac{1}{20}\sum_{i=1}^{n} t_i = 2,$$
$$\mathrm{var}(t) = \frac{20}{19}\left[\frac{1}{20}\sum_{i=1}^{n}(t_i - 2)^2\right] = .947368.$$

The two sets of values has to be equal since the basis of EDF representation too is that each of the observed outcomes have equal probability.

**In the next four questions below, use the following *iid* random variable outcomes from a** $Uniform(0,1)$ **probability distribution:**

| | | | | |
|---|---|---|---|---|
| **0.2957** | **0.3566** | **0.8495** | **0.5281** | **0.0914** |
| **0.5980** | **0.4194** | **0.9722** | **0.7313** | **0.1020** |
| **0.5151** | **0.6369** | **0.7888** | **0.9893** | **0.1252** |
| **0.6362** | **0.1392** | **0.1510** | **0.4202** | **0.2946** |
| **0.1493** | **0.0565** | **0.4959** | **0.8899** | **0.6343** |

29. a)

| **P(X)** | **X=-2.5log(1-P)** | **P(X)** | **X=-2.5log(1-P)** |
|---|---|---|---|
| 0.2957 | 0.8764 | 0.9893 | 11.3438 |
| 0.3566 | 1.1025 | 0.1252 | 0.3344 |
| 0.8495 | 4.7345 | 0.6362 | 2.5279 |
| 0.5281 | 1.8775 | 0.1392 | 0.3747 |
| 0.0914 | 0.2396 | 0.1510 | 0.4092 |
| 0.5980 | 2.2783 | 0.4202 | 1.3627 |
| 0.4194 | 1.3592 | 0.2946 | 0.8725 |
| 0.9722 | 8.9568 | 0.1493 | 0.4042 |
| 0.7313 | 3.2854 | 0.0565 | 0.1454 |
| 0.1020 | 0.2690 | 0.4959 | 1.7125 |
| 0.5151 | 1.8095 | 0.8899 | 5.5159 |
| 0.6369 | 2.5327 | 0.6343 | 2.5149 |
| 0.7888 | 3.8874 | | |

b) Sample mean = 2.5545, Sample variance = 5.5065, Estimated mean = 2.5545, Estimated variance = 5.7359 True mean =2.5, True variance = 6.25.

31. a)

| P(X) | X=5(1−P)^(−.5) | P(X) | X=5(1−P)^(−.5) |
|---|---|---|---|
| 0.2957 | 5.9579 | 0.9893 | 48.3368 |
| 0.3566 | 6.2335 | 0.1252 | 5.3458 |
| 0.8495 | 12.8885 | 0.6362 | 8.2897 |
| 0.5281 | 7.2786 | 0.1392 | 5.3891 |
| 0.0914 | 5.2455 | 0.1510 | 5.4265 |
| 0.5980 | 7.8860 | 0.4202 | 6.5665 |
| 0.4194 | 6.5619 | 0.2946 | 5.9532 |
| 0.9722 | 29.9880 | 0.1493 | 5.4210 |
| 0.7313 | 9.6458 | 0.0565 | 5.1475 |
| 0.1020 | 5.2763 | 0.4959 | 7.0423 |
| 0.5151 | 7.1803 | 0.8899 | 15.0687 |
| 0.6369 | 8.2977 | 0.6343 | 8.2681 |
| 0.7888 | 10.8799 | | |

b) Sample mean = 9.9830, Sample variance = 90.2225, Estimated mean = 9.9830, Estimated variance = 93.9817 True mean =10, True variance is not defined.

33. a) In order to apply the multivariate change of variables approach, define $\begin{bmatrix} r \\ w \end{bmatrix} = \begin{bmatrix} pq \\ p \end{bmatrix}$, which has the inverse $\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} w^{-1}r \\ w \end{bmatrix}$. The Jacobian of the inverse function system is given by $J = \begin{bmatrix} w^{-1} & -rw^{-2} \\ 0 & 1 \end{bmatrix}$ which has the determinant $|J| = w^{-1}$. Then the joint density of $r$ and $w$ is given by $f(r,w) = .01e^{-.01r} I_{[1.50,2.50]}(w) I_{[0,\infty)}(rw^{-1})$. The marginal density of $r$ *is then given by* $f(r) = \int_{1.50}^{2.50} .01e^{-.01r} I_{[1.50,2.50]}(w) I_{[0,\infty)}(r)dw = .01e^{-.01r} I_{[0,\infty)}(r)$.

b) $\mathrm{E}(R) = \int_0^\infty r\ .01e^{-.01r} dr = 100 \Rightarrow \$100,000$.

c) $P(r > 100) = \int_{100}^\infty .01e^{-.01r} dr = .3678$.

**Chapter 7 – Student Answer Key – Odd Numbers Only**
**Elements of Point Estimation Theory**

1. a) The joint density is

$$X \sim f(\mathbf{x};\theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta} \prod_{i=1}^n I_{(0,\infty)}(x_i),$$

which belongs to the exponential class of density functions (Theorem 7.9),

$$f(\mathbf{x};\theta) = \exp\left(c(\theta)g(\mathbf{x}) + d(\theta) + z(\mathbf{x})\right) I_A(\mathbf{x}),$$

where $c(\theta) = -\frac{1}{\theta}, g(\mathbf{x}) = \sum_{i=1}^n x_i, z(\mathbf{x}) = 0, d(\theta) = -\ln(\theta^n), A = \times_{i=1}^n (0,\infty)$. By Theorem 7.4, $g(\mathbf{x}) = \sum_{i=1}^n X_i$ is a minimal sufficient statistic.

b) Note that $c(\theta) = -1/\theta,\ \forall\theta > 0,$ has a range of $(-\infty,0)$ containing an open interval. By Theorem 7.8, $g(\mathbf{x}) = \sum_{i=1}^n X_i$ is a complete sufficient statistic.

c) By Theorem 7.10, $\bar{X} = \sum_{i=1}^n X_i / n$ is a complete sufficient statistic. In addition, $\mathrm{E}(\bar{X}) = \theta$. Thus, Lehmann-Scheffé's Completeness Theorem (Theorem 7.22) implies $\bar{X}$ is the MVUE of $\theta$. Here, the estimate of $\theta$ is $\bar{x} = 28.7$.

d) Theorem 7.5 (Rao-Blackwell) implies one need only consider functions of sufficient statistics to define an MVUE for $\theta^2$. Examine

$$\left(g(\mathbf{X})\right)^2 = \left(\sum_{i=1}^n X_i\right)^2 = \sum_{i=1}^n X_i^2 + 2\sum\sum_{i<j} X_i X_j,$$ which has expectation

$$\mathrm{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] = \sum_{i=1}^n \mathrm{E}X_i^2 + 2\sum\sum_{i<j} \mathrm{E}X_i \mathrm{E}X_j \text{ (by independence)}$$

$$= n\left(2\theta^2\right) + n(n-1)\theta^2 = n(n+1)\theta^2$$

Defining $t(\mathbf{X}) = \frac{\left(\sum_{i=1}^n X_i\right)^2}{n(n+1)}$ yields $\mathrm{E}(t(\mathbf{X})) = \theta^2$. So by the Lehmann-Scheffé Completeness Theorem and Theorem 7.10, $t(\mathbf{X})$ is the MVUE for $\theta^2$. The estimate for $\theta^2$ is $t(\mathbf{x}) = (n\bar{x})^2 / (n(n+1)) = (200(28.7))^2 / (200(201)) = 819.59$.

e) Again, by the Lehmann-Scheffé Completeness Theorem and Theorem 7.10

$$t(\mathbf{X}) = \frac{2\left(\sum_{i=1}^{n} X_i\right)^2}{n(n+1)},$$

which has E($t$(**X**)) = $2\theta^2$, is the MVUE for $2\theta^2$. The estimate for $2\theta^2$ is 1639.18

f) Using the results above, and considering Theorem 7.12, define the vector estimator

$$\begin{bmatrix} \sum_{i=1}^{n} X_i / n \\ \left(\sum_{i=1}^{n} X_i\right)^2 \Big/ n(n+1) \\ 2\left(\sum_{i=1}^{n} X_i\right)^2 \Big/ n(n+1) \end{bmatrix}$$ which is the MVUE for $\begin{bmatrix} \theta \\ \theta^2 \\ 2\theta^2 \end{bmatrix}$ and produces estimates as

$$\begin{bmatrix} 28.7 \\ 819.59 \\ 1639.18 \end{bmatrix}.$$

g) No. Because a MVUE is unique with probability 1, and from part e) the MVUE for $2\theta^2$ is $2\left(\sum_{i=1}^{n} X_i\right)^2 \Big/ n(n+1)$.

h) No. The MVUE for $\theta^2$ is given in part d) by $\left(\sum_{i=1}^{n} X_i\right)^2 \Big/ n(n+1)$, which is unique with probability 1.

i)
- Yes, because it satisfies the conditions of the Lehmann-Scheffé Completeness Theorem. That is, $t$(**X**) is a function of a complete sufficient statistic $\left(\sum_{i=1}^{n} X_i\right)$ and is unbiased. Hence, $t$(**X**) is the MVUE for $P(z \le b) = 1 - e^{-b/\theta}$.
- Then $\hat{P}(z \le 20) = 1 - \left(1 - \frac{20}{(28.7)200}\right)^{200-1} I_{[20,\infty)}\left((28.7)200\right) = .5007.$

j)
- No, $t_*(\mathbf{X}) = 1 - e^{-b/\bar{X}}$ is not an MVUE for $F(b)$. In part i) it was shown $t$(**X**) was <u>the</u> unique MVUE.
- Consider $\operatorname{plim} t_*(\mathbf{X}) = 1 - e^{-b/\operatorname{plim}\bar{X}} = 1 - e^{-b/\theta} = F(b), \forall \theta > 0 \Rightarrow t_*(\mathbf{X})$ is consistent.

3.

| **Part** | **R(c)** | **Complete Sufficient Statistics** |
|---|---|---|
| a) | $(-\infty,0)\times(-\infty,\infty)$ | $\sum_{i=1}^{n}(\ln X_i)^2, \sum_{i=1}^{n}\ln X_i$ |
| b) | $(-1,\infty)$ | $\sum_{i=1}^{n}\ln X_i$ |
| c) | $(-\infty,\infty)$ | $\sum_{i=1}^{n}X_i$ |
| d) | $(-\infty,0)$ | $\sum_{i=1}^{n}X_i$ |
| e) | $(-\infty,0)\times(-\infty,\infty)$ | $\sum_{i=1}^{n}X_i^2, \sum_{i=1}^{n}X_i$ |
| f) | $(-1,\infty)\times(-1,\infty)$ | $\sum_{i=1}^{n}\ln(X_i), \sum_{i=1}^{n}\ln(1-X_i)$ |

The joint densities in parts a), b), c), d), e), g) are all in the exponential class of densities (see Problem #2), and each $R(c)$ contains open rectangles. Hence, by Theorem 7.8, the minimal sufficient statistics defined in parts a), b), c), d), e), g) in Problem 2 are complete.

g) Completeness is shown through a verification of Definition 7.20. From 2f), $s(X)$ = max$\{X_1,...,X_n)$ is a minimal sufficient statistic. Let $h$ be any real valued function of $S$, and note that the expectation of $S$, set to zero $\forall\Theta\epsilon\Omega$, is defined by

$$\mathrm{E}h(S)=\int_0^{\Theta}h(s)\frac{ns^{n-1}}{\Theta^n}ds=0,\ \forall\Theta\in\Omega.$$

Note that $S$ is the $n^{\text{th}}$ order statistic, and so its pdf was obtained using Theorem 6.24 as

$$P(s\le b)=F(b)^n\Rightarrow f_s(b)=\frac{dF(b)^n}{db}=nF(b)^{n-1}f(b),$$

where $F(b)$ and $f(b)$ denote the CDF and PDF of $X_i$, and $f_s(b)$ denotes the PDF of $S$. Thus

$$f_s(b)=n\left(\frac{b}{\Theta}\right)^{n-1}\frac{1}{\Theta}I_{[0,\Theta]}(b)=n\frac{b^{n-1}}{\Theta^n}.$$

Since $\Theta>0$, it is clear that $\mathrm{E}h(S)=0\ \forall\Theta\epsilon\Omega$ iff

$$\int_0^{\Theta}h(s)ns^{n-1}ds=0\ \forall\Theta\in\Omega.$$

Using Leibniz's rule (Lemma 6.1, p. 332), this implies

$$\frac{d\int_0^{\Theta} h(s)\,ns^{n-1}ds}{d\Theta} = h(\Theta)\,n\Theta^{n-1} = 0,\ \forall\Theta\in\Omega \Rightarrow h(\Theta)=0\ \forall\Theta\in\Omega.$$

Thus, $s(X)$ is complete.

5. a) The $X_i$'s are iid by assumption, so that

$$f(x_1,\ldots,x_n;\lambda) = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}\prod_{i=1}^n I_{\{0,1,2,\ldots\}}(x_i);$$

which is in the exponential class since

$$f(x_1,\ldots,x_n;\lambda) = \exp\{c(\lambda)g(\mathbf{x}) + d(\lambda) + z(\mathbf{x})\}I_A(x),$$

where $c(\lambda) = \ln\lambda, g(\mathbf{x}) = \sum_{i=1}^n x_i, d(\lambda) = -n\lambda,$

$z(\mathbf{x}) = -\ln\left(\prod_{i=1}^n x_i!\right)$, and $A = \times_{i=1}^n \{0,1,2,\ldots\},$

which does not depend on $\lambda$.

Notice $c(\lambda)$ and $d(\lambda)$ are both <u>twice</u> continuously differentiable and $dc(\lambda)/d\lambda = 1/\lambda$, which has full rank since $\Omega = \{\lambda: \lambda > 0\}$ implies $\lambda\neq 0$, so that Theorem 7.16 applies and the CRLB regularity conditions are satisfied. Also notice $\Omega$ is defined to be an open interval on the real line, i.e., a 1-dimensional open rectangle.

b) Letting $T = t(\mathbf{X})$ be our unbiased estimator of $q(\lambda) = \lambda$, Corollary 7.4 gives the CRLB as

$$\operatorname{var}(t(X)) \geq \frac{1}{\mathrm{E}\left[\left(\frac{d\ln f(X_1,\ldots,X_n;\lambda)}{d\lambda}\right)^2\right]}.$$

Since we are in the exponential class, and $c(\lambda)$ and $d(\lambda)$ are both twice continuously differentiable we have that (see p. 416)

$$\mathrm{E}\left[\left(\frac{d\ln f(X_1,\ldots,X_n;\lambda)}{d\lambda}\right)^2\right] = -\mathrm{E}\left[\frac{d^2\ln f(X_1,\ldots,X_n;\lambda)}{d\lambda^2}\right].$$

Now, $\ln f\left(x_1,\ldots,x_n;\lambda\right)=\left(\ln\lambda\right)\left(\sum_{i=1}^n x_i\right)-n\lambda-\ln\left(\prod_{i=1}^n x_i!\right)$ (suppressing the indicator function); from which,

$$\left[\frac{d^2\ln f\left(\mathbf{x};\lambda\right)}{d\lambda^2}\right]=-\left(\sum_{i=1}^n x_i\right)/\lambda^2, \text{ hence } -\mathrm{E}\left[\frac{d^2\ln f\left(\mathbf{X};\lambda\right)}{d\lambda^2}\right]=-\left(-\frac{1}{\lambda^2}\sum_{i=1}^n \mathrm{E}X_i\right)=n/\lambda.$$

The CRLB is then expressed as $\operatorname{var}\left(t\left(\mathbf{X}\right)\right)\geq\lambda/n$. $\bar{X}$ is the MVUE for estimating $\lambda$. This follows since

$\mathrm{E}\left(\bar{X}\right)=n^{-1}\sum_{i=1}^n \mathrm{E}X_i=\lambda$ (i.e., $\bar{X}$ is unbiased), and $\operatorname{var}\left(\bar{X}\right)=n^{-2}\operatorname{var}\left(\sum_{i=1}^n X_i\right)=\lambda/n$, which is equal to the CRLB.

c) From Theorem 7.20 (CRLB Attainment), define the following

$$t\left(\mathbf{X}\right)=q\left(\lambda\right)+\frac{dq\left(\lambda\right)}{d\lambda}\left[-\mathrm{E}\left(\frac{d^2\ln f\left(\mathbf{X};\lambda\right)}{d\lambda^2}\right)\right]^{-1}\frac{d\ln f\left(\mathbf{x};\lambda\right)}{d\lambda},$$

$$=\lambda+1\left[n/\lambda\right]^{-1}\left[\left(1/\lambda\right)\sum_{i=1}^n x_i-n\right]=\lambda+\frac{1}{n}\left[n\left(\frac{\bar{x}}{\lambda}-1\right)\right]=\lambda+\bar{x}-\lambda=\bar{x}.$$

Since $\bar{X}=n^{-1}\sum_{i=1}^n X_i$ is a statistic (i.e., it does not depend upon $\lambda$), Theorem 7.20 implies that it is the MVUE for estimating $\lambda$. Given $n$=100 and $\sum_{i=1}^{100} x_i=283$, $\hat{\lambda}=\bar{x}=2.83$.

d) From LLCLT and Slutsky's Theorems,

$$n^{1/2}\left(\bar{X}-\lambda\right)\xrightarrow{\mathrm{d}}N\left(0,\lambda\right),$$

and since λ>0 (by the parameterization of the Poisson density) implies a positive definite (1×1) covariance matrix; we have that $\bar{X}$ is a CAN estimator (see page 388).

It follows that

$$\bar{X}_n\overset{a}{\sim}N\left(\lambda,\lambda/n\right),$$

from which we see the asymptotic covariance of $\bar{X}_n$ is equal to the CRLB. Thus $\bar{X}_n$ is asymptotically efficient (see p. 391 and p. 422).

e) Letting $t_1(\mathbf{X})$ be our unbiased estimator of $e^{-\lambda}$, Corollary 7.3 implies

$$\operatorname{var}\left(t_1\left(\mathbf{X}\right)\right)\geq\left[\frac{d}{d\lambda}\left(e^{-\lambda}\right)\right]^2\frac{\lambda}{n}=\frac{e^{-2\lambda}\lambda}{n}.$$

An unbiased estimator of $e^{-\lambda}$ which achieves the CRLB does not exist. To see this, apply the CRLB Attainment Theorem (Theorem 7.20):

$$t_1(\mathbf{X}) = q(\lambda) + \frac{dq(\lambda)}{d\lambda}\left[\mathrm{E}\left(\left[\frac{d\ln f(\mathbf{X},\lambda)}{d\lambda}\right]^2\right)\right]^{-1}\frac{d\ln f(\mathbf{x};\lambda)}{d\lambda}$$

$$= e^{-\lambda} + \left(-e^{-\lambda}\right)\left(\frac{\lambda}{n}\right)\left[\frac{\sum_{i=1}^n x_i}{\lambda} - n\right] = e^{-\lambda}\left[1-(\bar{x}-\lambda)\right].$$

Since this does not define a statistic, Theorem 7.20 implies the nonexistence of an unbiased estimator which achieves the CRLB.

7. a) • The alternative estimator is not unbiased, since

$$\mathrm{E}t_n(\mathbf{X}) = \mathrm{E}\left(\sum_{i=1}^n X_i/(n+k)\right) = \frac{n\mu}{n+k} \neq \mu \text{ when } k \neq 0.$$

• It is asymptotically unbiased, since

$$\lim_{n\to\infty}\mathrm{E}t_n(\mathbf{X}) = \lim_{n\to\infty}\left(\frac{n\mu}{n+k}\right) = \lim_{n\to\infty}\left(\frac{n}{n+k}\right)\lim_{n\to\infty}(\mu) = \mu,$$

$t_n(\mathbf{X})$ is in the CAN class, and the second order moments of $t_n(\mathbf{X})$ are bounded (see p. 387). An alternative proof is given in b) below.

• It is not BLUE, since it is not unbiased.

• It is consistent, since

$$\text{plim } t_n(\mathbf{X}) = \text{plim}\left(\sum_{i=1}^n X_i/(n+K)\right) = \text{plim}\left[\left(\frac{n}{n+k}\right)\bar{X}_n\right] = \text{ plim}\left(\frac{n}{n+k}\right)\text{ plim }\bar{X}_n = \mu.$$

b) We have already proved the general result that when random sampling is from a population distribution,

$$\bar{X}_n \overset{a}{\sim} N\left(\mu,\sigma^2/n\right).$$

Regarding the estimator $t_n(\mathbf{X})$, note that

$$n^{1/2}\left(\frac{1}{n+k}\sum_{i=1}^n X_i - \mu\right) = n^{1/2}\left(\frac{n}{n+k}\bar{X}_n - \mu\right) = n^{1/2}\left(\bar{X}_n-\mu\right) - \frac{n^{1/2}k}{n+k}\bar{X}_n \xrightarrow{d} N\left(0,\sigma^2\right),$$

by Slutsky's theorem, since $\bar{X}_n \xrightarrow{p} \mu$ and $n^{1/2}k/_{n+k} \to 0,$ so that $t_n(\mathbf{X}) \overset{a}{\sim} N(\mu,\sigma^2).$ Thus $t_n(\mathbf{X})$ is in the CAN class, is asymptotically unbiased, and is not distinguishable from $\bar{X}_n$ on the basis of the respective asymptotic distributions.

c)

$$\mathrm{MSE}(\bar{X}_n) = \mathrm{var}(\bar{X}_n) + \left[\mathrm{bias}(\bar{X}_n)\right]^2 = \sigma^2 / n,$$

$$\begin{aligned}\mathrm{MSE}(t_n(\mathbf{X})) &= \mathrm{var}(t_n(\mathbf{X})) + \left[\mathrm{bias}(t_n(\mathbf{X}))\right]^2 \\ &= \frac{n}{(n+k)^2}\sigma^2 + \frac{k^2}{(n+k)^2}\mu^2.\end{aligned}$$

Consider the relative efficiency of the two estimators

$$\frac{\mathrm{MSE}(t_n(\mathbf{X}))}{\mathrm{MSE}(\bar{X}_n)} = \frac{n^2}{(n+k)^2} + \frac{nk^2}{(n+k)^2}\left(\frac{\mu^2}{\sigma^2}\right).$$

Note, MSE($t_n(\mathbf{X})$) < MSE ( $\bar{X}_n$ ) iff

$$\frac{n^2}{(n+k)^2} + \frac{nk^2}{(n+k)^2}\left(\frac{\mu^2}{\sigma^2}\right) < 1, \text{ or alternatively iff } \frac{\mu^2}{\sigma^2} < \frac{2n+k}{nk},$$

implying $t_n(\mathbf{X})$ will be relatively more efficient (better in terms of MSE) than $\bar{X}_n$ if the above inequality is met. Depending on values of $(\mu^2/\sigma^2)$, ∃ choices of $k$ that satisfies the inequality.

d) Since we do not know the value of $(\mu^2/\sigma^2)$, we do not know which choices of $k$ would result in a MSE improvement.

9. a) The random sample can be thought of as the outcomes of 10,000 iid random variables, each being geometrically distributed. The statistical model for the sampling experiment is given by

$$f(\mathbf{x};p) = p^n (1-p)^{\left(\sum_{i=1}^n x_i - n\right)} \prod_{i=1}^n I_{\{1,2,3,\ldots\}}(x_i),$$

and $p \in \Omega = (0,1)$. $X_i$ is the number of trials necessary to get the first failure for the $i^{\text{th}}$ trigger.

b) Yes. $\mathrm{E}\bar{X}_n = n^{-1}\sum_{i=1}^n \mathrm{E}X_i = n^{-1}(n\mu) = \mu = 1/p,\ \forall p \in \Omega.$

c) Yes. $\lim_{n\to\infty} \mathrm{E}\bar{X}_n = 1/p,\ \forall p \in \Omega,$ and $\bar{X}_n$ is in the CAN class with bounded second moments (p. 387). Alternative proof in i) below.

d) Yes. By Khinchin's WLLN, $\bar{X}_n \xrightarrow{\mathrm{p}} \mu = 1/p.$

e) Let $r_0 = 1$ in problem 2d), then $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic. By Lehmann-Scheffé's Completeness Theorem, $\bar{X}_n$ is the MVUE for $\mu$. Hence, $\bar{X}$ is BLUE. Alternatively, the result of Example 7.6, p. 384, can be used.

f) The CRLB is

$$\left[\frac{dq(p)}{dp}\right]^2 \left[-\mathrm{E}\frac{d^2 \ln f(\mathbf{X};p)}{dp^2}\right]^{-1} = \frac{1-p}{np^2},$$

where $\ln f(\mathbf{x};p) = n\ln(p) + \left(\sum_{i=1}^{n} x_i - n\right)\ln(1-p) + \ln\left[\prod_{i=1}^{n} I_{\{1,2,3,\ldots\}}(x_i)\right]$

$$\Rightarrow \frac{d\ln f(\mathbf{x};p)}{dp} = \frac{n}{p} - \frac{\left(\sum_{i=1}^{n} x_i - n\right)}{1-p}$$

$$\Rightarrow \frac{d^2 \ln f(\mathbf{x};p)}{dp^2} = -\frac{n}{p^2} - \frac{\left(\sum_{i=1}^{n} x_i - n\right)}{(1-p)^2}$$

$$\Rightarrow \mathrm{E}\left[\frac{d^2 \ln f(\mathbf{X};p)}{dp^2}\right] = -\frac{n}{p^2} - \frac{(n/p - n)}{(1-p)^2} = \frac{-n}{p^2(1-p)} \text{ and } \frac{dq(p)}{dp} = \frac{-1}{p^2}.$$

g) Yes. See part e).

h) The CRLB Attainment Theorem requires that

$$t(\mathbf{x}) = q(p) + \frac{dq(p)}{dp}\left[-\mathrm{E}\frac{d^2 \ln f(\mathbf{X};p)}{dp^2}\right]^{-1} \frac{d\ln f(\mathbf{x};p)}{dp}$$

$$= \frac{1}{p} + \left(\frac{-1}{p^2}\right)\left[\frac{p^2(1-p)}{n}\right]\left(\frac{n}{p} - \frac{\left(\sum_{i=1}^{n} x_i - n\right)}{1-p}\right)$$

$$= \frac{1}{p} - \left(\frac{1-p}{n}\right)\left(\frac{n}{p} - \frac{\sum_{i=1}^{n} x_i}{1-p} + \frac{n}{1-p}\right) = \bar{x}_n.$$

Hence, $\bar{X}_n$ achieves the CRLB and is MVUE for $1/p$.

i) By the LLCLT, and Definition 5.2 $\bar{X}_n \overset{a}{\sim} N\left(\frac{1}{p}, n^{-1}\left(\frac{1-p}{p^2}\right)\right)$.

Since the asymptotic covariance matrix is equal to the CRLB, $\bar{X}_n$ is an asymptotically efficient estimator. The estimator is in the CAN class, and is asymptotically unbiased for estimating $1/p$.

j) $\bar{x}_n = \sum_{i=1}^{10,000} x_i / 10,000 = 1500.$

k) Yes, since $\text{plim}\left(\bar{X}_n^{-1}\right) = \frac{1}{\text{plim}\,\bar{X}_n} = p,\ \forall p \in (0,1)$.

l) Referring to Theorem 5.39, letting $g(\mathbf{X}) = \left(\bar{X}_n\right)^{-1}$, then

$$\left(\bar{X}_n\right)^{-1} \overset{a}{\sim} N\left(g(\mu), n^{-1} G \Sigma G'\right),$$

where $g(\mu) = p, G = \left.\frac{dg(\bar{x}_n)}{d\bar{x}_n}\right|_{\bar{x}_n=\mu} = -\left(\bar{x}_n\right)^{-2}\Big|_{\bar{x}_n=\mu} = -p^2$, and $\Sigma = \sigma^2 = \frac{(1-p)}{p^2}$

$$\Rightarrow \left(\bar{X}_n\right)^{-1} \overset{a}{\sim} N\left(p, \frac{p^2(1-p)}{n}\right).$$

m) First, $\left(\bar{x}_n\right)^{-1} = (1500)^{-1} = .00067$ is an estimate of $p$. Second, from part i) $\sigma^2$ = (1 - .00067)/[(10,000)(.00067)$^2$] = 224.85. Then

$$P\left(\left|\bar{x}_n - \frac{1}{p}\right| \le 50\right) = P\left(\left|\frac{\bar{x}_n - 1/p}{14.995}\right| \le \frac{50}{14.995}\right) = P\left(|z| \le 3.334\right) = .9996.$$

n) The expected number of impacts before failure is 1500, which surpasses the requirement of 1000. Also, the outcomes of $\bar{X}_n$ are within 50 units of the true number of impacts with probability .9996. Hence, one could say with very high confidence that the trigger mechanisms meet the Detroit Manufacturer's requirements.

11. a)

$$f\left(y_1,\ldots,y_n;\mu,\sigma^2\right)=\frac{1}{(2\pi)^{n/2}\left(\sigma^n\right)\left(\prod_{i=1}^n y_i\right)}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n\left(\ln y_i-\mu\right)^2\right]\prod_{i=1}^n I_{(0,\infty)}\left(y_i\right)$$

$$=\exp\left[\sum_{i=1}^2 c_i\left(\mu,\sigma^2\right)g_i(\mathbf{y})+d\left(\mu,\sigma^2\right)+z(\mathbf{y})\right]I_A(\mathbf{y}),$$

where

$$c_1\left(\mu,\sigma^2\right)=-\frac{1}{2\sigma^2},g_1(\mathbf{y})=\sum_{i=1}^n\left(\ln y_i\right)^2,c_2\left(\mu,\sigma^2\right)=\frac{\mu}{\sigma^2},g_2(\mathbf{y})=\sum_{i=1}^n \ln y_i$$

$$d\left(\mu,\sigma^2\right)=-\frac{n\mu^2}{2\sigma^2}-n\ln\left((2\pi)^{1/2}\sigma\right),z(\mathbf{y})=-\ln\left[\prod_{i=1}^n y_i\right],A=\underset{i=1}{\overset{n}{\times}}(0,\infty).$$

Hence, $f(y_1,\ldots,y_n;\mu,\sigma^2)$ is in the exponential class of densities, and $\sum_{i=1}^n\left(\ln Y_i\right)^2,\sum_{i=1}^n \ln Y_i$ are minimal sufficient statistics since $c_1$ and $c_2$ are linearly independent.

b) Yes. Since $R(c)=(-\infty,0)\times(-\infty,\infty)$ contains an open rectangle they are complete sufficient statistics.

c)
- Define $t_i(\mathbf{Y})=n^{-1}\sum_{i=1}^n \ln Y_i$ and note that $\mathrm{E}\left(t_1(\mathbf{Y})\right)=n^{-1}\sum_{i=1}^n \mathrm{E}\left(\ln Y_i\right)=n^{-1}(n\mu)=\mu.$ Since $t_1(\mathbf{Y})$ is unbiased and a function of complete sufficient statistics, by Lehmann-Scheffé's Completeness Theorem, $t_1(\mathbf{Y})$ is the MVUE for $\mu$.

- $\operatorname{var}\left(t_1(\mathbf{Y})\right)=n^{-2}\sum_{i=1}^n \operatorname{var}\left(\ln \mathbf{Y}_i\right)=\dfrac{\sigma^2}{n}\Rightarrow\lim_{n\to\infty}\operatorname{var}\left(t_1(\mathbf{Y})\right)=0\Rightarrow t_1(\mathbf{Y})\xrightarrow{m}\mu\Rightarrow t_1(\mathbf{Y})\xrightarrow{p}\mu.$
  Yes, $t_1(\mathbf{Y})$ is consistent.

d)
- The random variables $\ln\mathbf{Y}_1,\ldots,\ln\mathbf{Y}_n$ are iid $N(\mu,\sigma^2)$, so that $\mathrm{E}(t_2(\mathbf{Y}))=\sigma^2$. Also, $t_2(\mathbf{Y})$ is a function of complete sufficient statistics. By Lehmann-Scheffé's Completeness Theorem, $t_2(\mathbf{Y})$ is the MVUE for $\sigma^2$.

- $\operatorname{var}\left(t_2(\mathbf{Y})\right)=n^{-1}\left(\mu_4-\left(\dfrac{n-3}{n-1}\right)\sigma^4\right)\to 0$ as $n\to\infty$ (see p. 317)

  $\Rightarrow t_2(\mathbf{Y})\xrightarrow{m}\sigma^2\Rightarrow t_2(\mathbf{Y})\xrightarrow{p}\sigma^2$. Yes, $t_2(\mathbf{Y})$ is consistent.

e) Let $t_3(\mathbf{Y})=e^{t_1(\mathbf{Y})+t_2(\mathbf{Y})/2}$, which is a continuous function. By properties of the plim operator,

$$\operatorname{plim} t_3(\mathbf{Y})=e^{\operatorname{plim} t_1(\mathbf{Y})+\operatorname{plim} t_2(\mathbf{Y})/2}=e^{\mu+\sigma^2/2}\text{ (from parts c,d)}$$

Hence, $t_3(\mathbf{Y})$ is consistent.

f) Define an estimator of $\begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$ by $\begin{bmatrix} t_1(\mathbf{Y}) \\ t_2(\mathbf{Y}) \end{bmatrix}$, which is MVUE by Theorem 7.12. An estimate is $\begin{bmatrix} 3.75 \\ .5 \end{bmatrix}$. Using $t_3(\mathbf{Y})$ as our consistent estimator, the estimate of $e^{\mu+\sigma^2/2}$, is given by $e^{3.75+.5/2} = 54.6$.

13.

$$\text{MSE } t(\mathbf{X}) = \text{var}\, t(\mathbf{X}) = \frac{1}{4n^2}\sum_{i=1}^{n} \text{var}\left(X_i^2\right) = \frac{1}{4n^2}\left[n\left[\text{E}X_i^4 - \left(\text{E}X_i^2\right)^2\right]\right] = \frac{1}{4n^2}\left[n24\theta^4 - n\left(2\theta^2\right)^2\right] = 5\theta^4 / n,$$

$$\text{MSE}\left(S^2\right) = \text{var}\left(S^2\right) + \left(\text{bias}\, S^2\right)^2 = \left[\left(\frac{n-1}{n}\right)^2 \mu_4 - \frac{(n-1)(n-3)}{n^2}\sigma^2\right] / n + \frac{\sigma^4}{n^2}$$

$$= \left[\left(\frac{n-1}{n}\right)^2 9\theta^4 - \frac{(n-1)(n-3)}{n^2}\theta^4\right] / n + \frac{\theta^4}{n^2} \geq \frac{(n-1)^2}{n^2}\frac{8\theta^4}{n}$$

$$\frac{\text{MSE } t(\mathbf{X})}{\text{MSE } S^2} \leq \frac{5}{8}\left(\frac{n}{n-1}\right)^2 \Rightarrow \text{ prefer } t(\mathbf{X}) \text{ whenever } \left(\frac{n}{n-1}\right)^2 \leq \frac{8}{5}.$$

15. a) The joint density of the random sample can be represented as

$$\left(\mathbf{X}_1,\ldots,\mathbf{X}_\mathbf{n}\right) \sim \frac{1}{(2\pi)^n \,|\,\Sigma\,|^{n/2}} \exp\left[-(1/2)\sum_{i=1}^{n}\left(\mathbf{x}_\mathbf{i} - \boldsymbol{\mu}\right)' \Sigma^{-1}\left(\mathbf{x}_\mathbf{i} - \boldsymbol{\mu}\right)\right],$$

where $x_i$ is a $2\times1$ vector of observed rates of returns on the two stocks in time period $i$, $\mu$ is a $2\times1$ vector of unknown population mean rates of return for the two stocks, and $\Sigma$ is the unknown population covariance matrix.

A strategy for finding the minimal, complete (vector) sufficient statistic involves showing that the joint density of the random sample is in the exponential class of densities. We already know the bivariate normal density is in the exponential class; so that, by Theorem 7.9, the joint density of the random sample will also be in the exponential class. Proceeding along these lines, the joint density of the random sample can be rewritten as

$$f\left(\mathbf{x}_1,\ldots,\mathbf{x}_\mathbf{n};\boldsymbol{\mu},\Sigma\right) = \frac{1}{(2\pi)^n \,|\,\Sigma\,|^{n/2}} \exp\left[-1/2\left(\sum_{i=1}^{n}\left(\mathbf{x}_\mathbf{i}'\Sigma^{-1}\mathbf{x}_\mathbf{i} - 2\boldsymbol{\mu}'\Sigma^{-1}\mathbf{x}_\mathbf{i} + \boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}\right)\right)\right].$$

Decomposing the expressions $\mathbf{x}_\mathbf{i}'\Sigma^{-1}\mathbf{x}_\mathbf{i}$ and $\boldsymbol{\mu}'\Sigma^{-1}\mathbf{x}_\mathbf{i}$ we may further rewrite the joint density as

$$f(\mathbf{x_1},\ldots,\mathbf{x_n};\boldsymbol{\mu},\Sigma)=\frac{1}{(2\pi)^n\,|\,\Sigma\,|^{n/2}}\exp\Bigg[-1/2\Bigg(\Sigma^{-1}_{(11)}\sum_{i=1}^n x_{i1}^2+\Sigma^{-1}_{(22)}\sum_{i=1}^n x_{i2}^2$$
$$+2\Sigma^{-1}_{(12)}\sum_{i=1}^n x_{i1}x_{i2}-2\left(\boldsymbol{\mu}_1\Sigma^{-1}_{(11)}+\boldsymbol{\mu}_2\Sigma^{-1}_{(21)}\right)\sum_{i=1}^n x_{i1}-2\left(\boldsymbol{\mu}_1\Sigma^{-1}_{(12)}+\boldsymbol{\mu}_2\Sigma^{-1}_{(22)}\right)\sum_{i=1}^n x_{i2}+n\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}\Bigg)\Bigg].$$

where $x_{ij}$, $i$=1,...,$n$, $j$=1,2 refers to the $i^{\text{th}}$ observed rate of return on stock $j$ and $\Sigma^{-1}_{(\ell,m)}$ refers to element ($\ell$, $m$) of $\Sigma^{-1}$.

It can now be seen that the joint density is of the exponential class form, i.e.,

$f(\mathbf{x}_1,\ldots,\mathbf{x_n};\,\boldsymbol{\mu},\Sigma)=\exp[c(\boldsymbol{\Theta})'g(\mathbf{x})+d(\boldsymbol{\Theta})+z(\mathbf{x})]$,

where

$$\underset{(1\times5)}{c(\Theta)'}=-\left[1/2\Sigma^{-1}_{(11)},1/2\Sigma^{-1}_{(22)},\Sigma^{-1}_{(12)},-\left(\mu_1\Sigma^{-1}_{(11)}+\mu_2\Sigma^{-1}_{(21)}\right),-\left(\mu_1\Sigma^{-1}_{(12)}+\mu_2\Sigma^{-1}_{(22)}\right)\right],$$
$$\underset{(1\times5)}{g(\mathbf{x})'}=\left[\sum_{i=1}^n x_{i1}^2,\sum_{i=1}^n x_{i2}^2,\sum_{i=1}^n x_{i1}x_{i2},\sum_{i=1}^n x_{i1},\sum_{i=1}^n x_{i2}\right],d(\theta)=-\frac{n}{2}\left[\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}+\ln\left[(2\pi)^2\,|\,\Sigma\,|\right]\right],z(\mathbf{x})=0$$

and the support of $\mathbf{X}$ does not depend upon $\boldsymbol{\mu}$ or $\Sigma$.

The range of each coordinate function of $c(\Theta)$ contains an open rectangle; thus, by Theorems 7.8 and 7.7, $g(\mathbf{X})$ is a minimal, complete vector sufficient statistic. Notice, we may consider the 5 elements of $g(\mathbf{X})$ as a set of minimal, complete sufficient statistics.

b) Consider $\bar{X}'=\left[n^{-1}\sum_{i=1}^n X_{i1},n^{-1}\sum_{i=1}^n X_{i2}\right]$ as a candidate; and notice $\mathrm{E}\bar{\mathbf{X}}'=[\mu_1,\mu_2]$. Since each coordinate function is a function of complete sufficient statistics, $\bar{\mathbf{X}}$ is the MVUE of $\mu$. (Lehmann-Scheffé's Completeness Theorem).

c) Consider as a candidate for the MVUE estimator of $\Sigma$, the matrix of sample variances and covariances, i.e.,

$$S=n^{-1}\begin{bmatrix}\sum_{i=1}^n\left(X_{i1}-\bar{X}_1\right)^2 & \sum_{i=1}^n\left(X_{i1}-\bar{X}_1\right)\left(X_{i2}-\bar{X}_2\right)\\ \sum_{i=1}^n\left(X_{i2}-\bar{X}_2\right)\left(X_{i1}-\bar{X}_1\right) & \sum_{i=1}^n\left(X_{i2}-\bar{X}_2\right)^2\end{bmatrix}.$$

To see clearly that the individual elements of $S$ are functions of complete sufficient statistics, we may rewrite $S$ as

$$S=n^{-1}\begin{bmatrix}\sum_{i=1}^n X_{i1}^2-n^{-1}\left(\sum_{i=1}^n X_{i1}\right)^2 & \sum_{i=1}^n X_{i1}X_{i2}-n^{-1}\left(\sum_{i=1}^n X_{i1}\right)\left(\sum_{i=1}^n X_{i2}\right)\\ \sum_{i=1}^n X_{i2}X_{i1}-n^{-1}\left(\sum_{i=1}^n X_{i2}\right)\left(\sum_{i=1}^n X_{i1}\right) & \sum_{i=1}^n X_{i2}^2-n^{-1}\left(\sum_{i=1}^n X_{i2}\right)^2\end{bmatrix}.$$

We know from the arguments presented in Chapter 6 (see especially Theorems 6.7 and 6.8) that the elements of the matrix above are biased estimators of the corresponding population variances and covariances in Σ. They become unbiased, however, if we change the scaling factor from $1/n$ to $1/(n\text{-}1)$. Thus,

$$\hat{\Sigma} = \left(\frac{n}{n-1}\right) S = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{21} & \hat{\sigma}_2^2 \end{bmatrix},$$

is the MVUE for Σ. It follows from Theorem 7.12, that $\left[\hat{\sigma}_1^2 \ \ \hat{\sigma}_2^2\right]'$ is the MVUE of $\left[\sigma_1^2 \ \ \sigma_2^2\right]'$, i.e., the diagonal of Σ.

d) Again by Theorem 7.12, the desired MVUE is $\left[\bar{X}_1, \bar{X}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2\right]'$, where the coordinates are defined as above.

e) Since $\hat{\sigma}_i^2 = \dfrac{nS_i^2}{n-1}, i = 1,2;$ our MVUE outcome is

[.048, .077, $.5102\times10^{-3}$, $.306\times10^{-4}$].

f) From Chapter 6, we know $\bar{X}_n \xrightarrow{p} \mu$ and $S_n^2 \xrightarrow{p} \sigma^2$; from which it follows

$$\text{plim } \hat{\sigma}_n^2 = \text{plim}\left(\frac{n}{n-1}\right) \text{plim}\left(S_n^2\right) = \sigma^2.$$

Since the plim is an element wise operator, it follows that

$$\text{plim}\left[\bar{x}_1, \bar{x}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2\right] = \left[\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\right],$$

i.e., the MVUE is a consistent estimator.

g) The expected dollar return is

$R = 500\mu_1 + 500\mu_2$.

By Theorem 7.12, our MVUE estimator of $R$ is

$$\hat{R} = 500\bar{X}_1 + 500\bar{X}_2$$

The outcome of which is \$62.50.

17. a) First estimator (sample mean) is unbiased, asymptotically unbiased, BLUE and consistent.

Second estimator:

$\mathrm{E}\left[t(X)\right] = np/(n+k) \neq p \Rightarrow$ not unbiased.
$\lim_{n\to\infty} \mathrm{E}\left[t(X)\right] = p \Rightarrow$ assymptotically unbiased.
Not the BLUE since not unbiased.
Consistent since $\lim_{n\to\infty} \operatorname{var}\left[t(X)\right] = 0 \Rightarrow \operatorname{plim}\left[t(X)\right] = p$.

b) $\bar{x}_n \overset{a}{\sim} N\left(p, \frac{p(1-p)}{n}\right), t(x) \overset{a}{\sim} N\left(\frac{np}{(n+k)}, \frac{p(1-p)}{(n+k)}\right)$.

Both are asymptotically unbiased and consistent but $\bar{x}_n$ has better finite sample properties.

c) $\mathrm{MSE}(\bar{x}_n) = \frac{p(1-p)}{n}, \mathrm{MSE}(t(x)) = \frac{p(1-p)}{(n+k)} + \frac{k^2p^2}{(n+k)^2}$.

The condition for $t(X)$ to be superior to $\bar{X}$ in terms of expected squared distance from *p:*

$$\frac{p(1-p)}{(n+k)} + \frac{k^2p^2}{(n+k)^2} \geq \frac{p(1-p)}{n}$$
$$\Rightarrow \frac{(1-p)}{(n+k)} + \frac{k^2p}{(n+k)^2} \geq \frac{(1-p)}{n}$$
$$\Rightarrow \frac{-(1-p)}{n} + \frac{kp}{(n+k)} \geq 0$$
$$\Rightarrow k \geq \frac{(1-p)(n+k)}{np} = \frac{(1-p)}{p} + \frac{(1-p)(k)}{np}$$
$$\Rightarrow k\left(1 - \frac{(1-p)}{np}\right) \geq \frac{(1-p)}{p}$$
$$\Rightarrow k \geq \frac{n(1-p)}{(np-1+p)}$$
$$\text{when} \lim_{n\to\infty} k \geq \frac{1}{p} - 1.$$

Accordingly, for an appropriate choice of *k,* $t(X)$ will be superior to $\bar{X}$ in terms of expected squared distance from *p.*

d) The best choice of value *k* is a function of unknown population mean *p*.

e) Assuming that the population mean *p* is equal to the sample mean $\bar{X}$, choose the optimum value of *k*.

19.

$$\mathrm{MSE}\left(\bar{X}_n\right) = \frac{\lambda}{n},$$

$$\mathrm{MSE}\left(\left(n-1\right)^{-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right) = n^{-1}\left[\left(\frac{n-1}{n}\right)^2 \lambda\left(1+3\lambda\right) - \frac{\left(n-1\right)\left(n-3\right)}{n^2}\lambda^2\right] > \frac{\lambda}{n}.$$

Accordingly, 3c) is a better estimator than 3d).

21. a) Assuming a Bernoulli distribution, the joint density function is

$$f\left(x;p\right) = \prod_{i=1}^{1000} p^{x_i}\left(1-p\right)^{1-x_i} I_{\{0,1\}}\left(x_i\right) = p^{\sum_{i=1}^{1000} x_i}\left(1-p\right)^{1000-\sum_{i=1}^{1000} x_i} \prod_{i=1}^{1000} I_{\{0,1\}}\left(x_i\right), \Omega = \left\{p : 0 \le p \le 1\right\}.$$

b) Define the sample mean $n^{-1}\sum_{i=1}^{n} X_i = \bar{X}$ as the MVLUE. Since $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for the Bernoulli distribution the linear estimator $\bar{X}$ is the MVUE by Lehmann-Scheffe completeness theorem. Therefore it is the MVLUE.

c) $\bar{X}$ is a consistent estimator since $\operatorname{plim}\left(\bar{X}\right) = p$. As $\mathrm{E}\left(\bar{X}\right) = p$ and $\operatorname{var}\left(\bar{X}\right) = {}^{p(1-p)}\!/_{n}$,

$$\bar{X} \overset{a}{\sim} = N\left(p, \frac{p\left(1-p\right)}{n}\right).$$

d) By the CRLB attainment theorem, the MVUE of the proportion of voters in favor of the initiative is $n^{-1}\sum_{i=1}^{n} X_i = \bar{X}$, which is the same as the linear estimator defined in a).

e)

Yes. It is given by

$$p^2 + \left(1-2p\right)n^{-1}\sum_{i=1}^{n} X_i = \hat{p}^2 + \left(1-2\hat{p}\right)n^{-1}\sum_{i=1}^{n} X_i$$

$$= \left(n^{-1}\sum_{i=1}^{n} X_i\right)^2 + \left(1-2n^{-1}\sum_{i=1}^{n} X_i\right)n^{-1}\sum_{i=1}^{n} X_i$$

$$= \left(1-n^{-1}\sum_{i=1}^{n} X_i\right)n^{-1}\sum_{i=1}^{n} X_i.$$

**Chapter 8 – Student Answer Key – Odd Numbers Only**
**Point Estimation Methods**

1. a) The transformed production function is

$$\ln Y_i = \beta_1^* + \beta_2 \ln \ell_i + \beta_3 \ln m_i + \varepsilon_i$$

where $\beta_1^* = \ln \beta_1$. Using the least-squares approach where

$\beta_* = \left[\beta_1^*, \beta_2, \beta_3\right]', \hat{\beta}_* = \left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}\mathbf{x}_*'\mathbf{Y}.$

$$\Rightarrow \mathbf{b}_* = \begin{bmatrix} .06156 & -.05020 & -.00177 \\ & .09709 & -.07284 \\ & \text{(symmetric)} & .11537 \end{bmatrix} \begin{bmatrix} 165.47200 \\ 180.32067 \\ 122.90436 \end{bmatrix},$$

$$\Rightarrow \mathbf{b}_* = \begin{bmatrix} .91720 \\ .24874 \\ .75100 \end{bmatrix}.$$

In addition,

$$\hat{S}_*^2 = \left(\mathbf{Y}_* - \mathbf{x}_*\hat{\beta}_*\right)'\left(\mathbf{Y}_* - \mathbf{x}_*\hat{\beta}_*\right)/(n-k) = \left(\mathbf{Y}_*'\mathbf{Y}_* - \mathbf{Y}_*'\mathbf{x}_*\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}\mathbf{x}_*'\mathbf{Y}_*\right)/(n-k),$$

$$\Rightarrow \hat{s}_*^2 = .2774 \times 10^{-3}.$$

b) (1-3) Since $\mathrm{E}\varepsilon = [\mathbf{0}]$, $\mathrm{E}\varepsilon\varepsilon' = \sigma^2\mathbf{I}$, and $\mathbf{x}_*$ has full column rank $\Rightarrow \hat{\beta}_*$ is unbiased and BLUE (by the Gauss-Markov Theorem). It is asymptotically unbiased assuming $n^{1/2}\left(\hat{\beta}_* - \beta_*\right) \xrightarrow{d} N\left([\mathbf{0}], \mathbf{Q}\right)$, where $\mathbf{Q}$ is a finite, positive definite matrix.

(4) Since $\varepsilon_i \sim$ iid $N(0,\sigma^2)$ and $\varepsilon$ is multivariate normally distributed $\Rightarrow \hat{\beta}_*$ is the MVUE (by Theorem 8.11).

(5) If $\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1} \to [\mathbf{0}]$ as $n \to \infty$, then $\hat{\beta}_* \xrightarrow{p} \beta_*$ (Theorem 8.4).

(6) Since $\hat{\beta}_* = \left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}\mathbf{x}_*'\mathbf{Y}_*$ is a linear combination of the elements of a multivariate normally distributed vector, $\Rightarrow \hat{\beta}_* \sim N\left(\beta_*, \sigma^2\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}\right)$.

c) (1-2) Since $\mathrm{E}\varepsilon = [\mathbf{0}]$, $\mathrm{E}\varepsilon\varepsilon' = \sigma^2\,\mathbf{I}$, and $\mathbf{x}$ has full column rank $\Rightarrow \hat{S}^2$ is unbiased. It will be asymptotically unbiased if $n^{1/2}\left(\hat{S}^2 - \sigma^2\right) \xrightarrow{d} N\left(0, \mu_4' - \sigma^4\right)$.

(3) $\hat{S}_1^2$ is not a linear estimator $\Rightarrow \hat{S}^2$ is not BLUE.

(4) Since $\varepsilon$ is multivariate normally distributed $\Rightarrow \hat{S}^2$ is MVUE (Theorem 8.11).

(5) Since the $\varepsilon_i$'s are iid $\Rightarrow \hat{S}^2$ is consistent for $\sigma^2$.

(6) Given ε is multivariate normally distributed $\Rightarrow \hat{S}^2$ has a gamma distribution (see p. 464).

d) By Theorem 7.12, if $\hat{\beta}_*$ is MVUE for $\beta_*$, then $\ell'\hat{\beta}_* = \hat{\beta}_2^* + \hat{\beta}_3^*$ is the MVUE for $\beta_2^* + \beta_3^*$, where $\ell' = [0\ 1\ 1]$.

- The estimated degree of homogeneity is

$$\hat{\beta}_2^* + \hat{\beta}_3^* = .24874 + .75100 = .99974.$$

- Since $\operatorname{plim}\left(\hat{\beta}_2^* + \hat{\beta}_3^*\right) = \operatorname{plim}\hat{\beta}_2^* + \operatorname{plim}\hat{\beta}_3^* = \beta_2^* + \beta_3^*$, the MVUE is consistent.

- The MVUE is normally distributed, since $\hat{\beta} \sim \mathrm{N}\left(\beta_*, \sigma^2\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}\right)$, and

$$\hat{\beta}_2^* + \hat{\beta}_3^* = \ell'\hat{\beta}_* \sim N\left(\ell'\beta_*, \sigma^2\ell'\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}\ell\right).$$

e) Since $\hat{S}^2$ is the MVUE for $\sigma^2$, and each entry in $\sigma^2\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}$ is a linear combination of $\sigma^2$, then by Theorem 7.12,

$\hat{S}^2\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}$ is the MVUE for $\sigma^2\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1}$.

The estimate is given by

$$\hat{s}^2\left(\mathbf{x}_*'\mathbf{x}_*\right)^{-1} = .2774\times10^{-3}\begin{bmatrix} .06156 & -.05020 & -.00177 \\ & .09709 & -.07284 \\ \text{(symmetric)} & & .11537 \end{bmatrix}.$$

3. a) $\hat{\beta} \sim N\left(\beta, \sigma^2\left(\mathbf{x}'\mathbf{x}\right)^{-1}\right)$, $(n-k)\hat{S}^2/\sigma^2 \sim \chi^2_{n-k}$ and $\hat{\beta}$ and $\hat{S}^2$ are independent (pp. 464-465).

Then $\ell'\left(\hat{\beta}-\beta\right)/\left(\sigma^2\ell'\left(\mathbf{x}'\mathbf{x}\right)^{-1}\ell\right)^{1/2} \sim N(0,1)$, so that

$$T = \frac{\ell'\left(\hat{\beta}-\beta\right)/\left(\sigma^2\ell'\left(\mathbf{x}'\mathbf{x}\right)^{-1}\ell\right)^{1/2}}{\left(\dfrac{(n-k)\hat{S}^2}{(n-k)\sigma^2}\right)^{1/2}} \sim t \text{ distribution with } (n-k) \text{ d.f.}$$

is the ratio of two independent random variables, the numerator being standard normal and the denominator being the square root of a $\chi^2$ random variable divided by its degrees of freedom.

b) From a) above,

P($-2.06 \leq t \leq 2.06$) = .95 (Table B.2)

$$\Rightarrow P\Big(\underbrace{\ell'\hat{\beta} - 2.06\left[s^2\ell'(\mathbf{x}'\mathbf{x})^{-1}\ell\right]^{1/2}}_{z_1} \leq \ell'\beta \leq \underbrace{\ell'\hat{\beta} + 2.06\left[s^2\ell'(\mathbf{x}'\mathbf{x})^{-1}\ell\right]^{1/2}}_{z_2}\Big] = .95.$$

5. a) Note that

$$f(\mathbf{x};\beta) = \beta^n e^{-\beta\sum_{i=1}^n x_i}\prod_{i=1}^n I_{(0,\infty)}(x_i) = \exp\left[n\ln\beta - \beta\sum_{i=1}^n x_i\right]\prod_{i=1}^n I_{(0,\infty)}(x_i),$$

so for the exponential class representation define

$$c(\beta) = \beta,\ g(\mathbf{x}) = \sum_{i=1}^n x_i,\ d(\beta) = n\ln\beta,\ z(\mathbf{x}) = 0,\ A = \mathop{\times}_{i=1}^n(0,\infty).$$

b) Because $R(c) = (0,\infty)$ contains an open interval, by Theorem 7.8 $\sum_{i=1}^n x_i$ is a complete sufficient statistic. Hence, it is minimal (Theorem 7.7).

c) Referring to Theorem 7.20 (attainment of the CRLB),

$$t(\mathbf{X}) = \beta + \frac{\beta^2}{n}\left[\frac{n}{\beta} - \sum_{i=1}^n X_i\right] = 2\beta - \beta^2\bar{X}_n,$$

which is not a statistic. Hence, there is no unbiased estimator that achieves the CRLB.

$$\left(\text{Here, } -\mathrm{E}\left[\frac{d^2\ln f(\mathbf{X};\beta)}{d\beta^2}\right] = \frac{n}{\beta^2} \text{ and } \frac{d\ln f(\mathbf{x};\beta)}{d\beta} = \frac{n}{\beta} - \sum_{i=1}^n x_i\right).$$

d) • The log-likelihood function is

$$\ln L(\beta;\mathbf{x}) = n\ln\beta - \beta\sum_{i=1}^n x_i + \ln\left(\prod_{i=1}^n I_{(0,\infty)}(x_i)\right).$$

The F.O.C. is

$$\frac{d\ln L(\beta;\mathbf{x})}{d\beta} = \frac{n}{\beta} - \sum_{i=1}^{n} x_i = 0.$$

$\Rightarrow \hat{\beta} = \left(\sum_{i=1}^{n} X_i / n\right)^{-1}$ is the ML estimator for $\beta$.

The second order condition is

$$\frac{d^2\ln L(\beta;\mathbf{x})}{d\beta^2} = \frac{-n}{\beta^2} < 0.$$

$\Rightarrow$ maximum at $\hat{\beta} = \left(\sum_{i=1}^{n} X_i / n\right)^{-1}$.

- Yes, $\hat{\beta}$ is a function of the complete sufficient statistic $\sum_{i=1}^{n} X_i$.
- Yes, $\hat{\beta} \xrightarrow{p} \beta$. Since, plim $\hat{\beta} = \frac{1}{\text{plim}\, \bar{X}_n} = \frac{1}{\theta} = \beta$.

e)
- Recall that $\beta = 1/\theta$ and that $\hat{\Theta} = \sum_{i=1}^{n} X_i / n$ is consistent (see Example 8.11), asymptotically normal, and asymptotically efficient (see Example 8.14) for estimating $\theta$. In addition, letting $q(\theta) = 1/\theta$, we have

$$\frac{dq(\theta)}{d\theta} = \frac{-1}{\theta^2}.$$

Hence, $q(\theta)$ is continuously differentiable $\forall\theta > 0$ and $dq(\theta)/d\theta$ is of full row rank. By Theorem 8.22,

$$q(\hat{\theta}) = \frac{1}{\left(\sum_{i=1}^{n} X_i / n\right)},$$

is a consistent, asymptotically normal, and asymptotically efficient MLE of $q(\theta) = 1/\theta = \beta$.

- $\hat{\beta} = \left(\sum_{i=1}^{n} X_i / n\right)^{-1}$ is not the MVUE of $\beta$. Since,

$$
\begin{aligned}
\mathrm{E}\left[\frac{n}{\sum_{i=1}^{n} X_i}\right] &= \mathrm{E}\left[\frac{n}{Y}\right] \quad \left(Y=\sum_{i=1}^{n} X_i \sim \text{Gamma}(n,\theta) \text{ from Theorem 4.2}\right) \\
&= n\int_0^{\infty}\left(\frac{1}{y}\right)\left(\frac{1}{\theta^n\Gamma(n)}y^{n-1}e^{-y/\theta}\right)dy \\
&= \frac{n}{\theta(n-1)}\underbrace{\int_0^{\infty}\frac{1}{\theta^{n-1}\Gamma(n-1)}y^{n-2}e^{-y/\theta}dy}_{\text{Gamma}(n-1,\theta)} \quad \left(\text{recall } \Gamma(n)=(n-1)\Gamma(n-1)\right) \\
&= \beta\left(\frac{n}{n-1}\right).
\end{aligned}
$$

Hence, $\hat{\beta}$ is a biased estimator of $\beta$, and thus cannot be MVUE.

7. a) It should be interpreted as a random sample generated from a composite experiment. Note that $\mathrm{E}Y_i = x_i\beta$. If $x_i$'s differ, then the expected values of the $Y_i$'s differ, and thus $\{Y_1,...,Y_n\}$ is a collection of independent (since $\varepsilon_i$'s are independent) but <u>not</u> identically distributed random variables.

b) $b=(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}=\dfrac{\sum_{i=1}^{n} x_y y_i}{\sum_{i=1}^{n} x_i^2}$ since $\mathbf{x}$ is a column vector. The least squares estimator is thus

$$\hat{\beta}=\frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}.$$

By the Gauss-Markov Theorem, we know $\hat{\beta}$ is BLUE. Without a distributional assumption such as normality on the error terms, however, we cannot demonstrate that $\hat{\beta}$ is MVUE.

c) Regarding the consistency of $\hat{\beta}$, note that

$$\text{var}\left(\hat{\beta}\right)=\frac{\sum_{i=1}^{n} x_i^2 \,\text{var}(Y_i)}{\left(\sum_{i=1}^{n} x_i^2\right)^2}=\frac{\sigma^2}{\sum_{i=1}^{n} x_i^2},$$

which follows directly from the fact the $\hat{\beta}$ is defined as a linear combination of the random variables $Y_1,...,Y_n$, and the $Y_i$'s have zero covariances. Since

$$\mathrm{E}\hat{\beta}=\frac{\sum_{i=1}^{n} x_i \mathrm{E}Y_i}{\sum_{i=1}^{n} x_i^2}=\frac{\sum_{i=1}^{n} x_i(x_i\beta)}{\sum_{i=1}^{n} x_i^2}=\beta,$$

if we additionally assume that $\sum_{i=1}^{n} x_i^2 \to \infty$ as $n \to \infty$, implying $\lim_{n\to\infty} \mathrm{var}\left(\hat{\beta}\right) = 0$, we have that $\hat{\beta} \xrightarrow{m} \beta$, and so $\hat{\beta} \xrightarrow{p} \beta$. Hence, $\hat{\beta}$ is consistent under this assumption. Regarding the asymptotic normality of $\hat{\beta}$, since $Y_i = x_i\beta + \varepsilon_i$, we may rewrite the least squares estimator as

$$\hat{\beta}_n = \frac{\sum_{i=1}^{n} x_i\left(x_i\beta + \varepsilon_i\right)}{\sum_{i=1}^{n} x_i^2} = \beta + \frac{\sum_{i=1}^{n} x_i\varepsilon_i}{\sum_{i=1}^{n} x_i^2},$$

hence,

$$\hat{\beta}_n - \beta = \frac{\sum_{i=1}^{n} x_i\varepsilon_i}{\sum_{i=1}^{n} x_i^2}.$$

It follows that

$$\frac{\hat{\beta}_n - \beta}{\left(\mathrm{var}\left(\hat{\beta}_n\right)\right)^{1/2}} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \frac{\left(\hat{\beta}_n - \beta\right)}{\sigma} = \sum_{i=1}^{n} \frac{x_i\varepsilon_i}{\left(\sigma^2 \sum_{i=1}^{n} x_i^2\right)^{1/2}} = \sum_{i=1}^{n} \frac{W_i}{\left(\sum_{i=1}^{n} \mathrm{var}\left(W_i\right)\right)^{1/2}},$$

where $W_i = x_i\varepsilon_i$. If we multiply the numerator and denominator by $1/n$, obtaining

$$\left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \frac{\left(\hat{\beta}_n - \beta\right)}{\sigma} = n^{1/2}\frac{\overline{W}_n}{\overline{\sigma}_n}, \text{ where } \overline{\sigma}_n = \left[n^{-1}\sum_{i=1}^{n} \mathrm{var}\left(W_i\right)\right]^{1/2},$$

then it is apparent that we may be able to apply Theorem 5.32 under certain assumptions.

First, notice $\mathrm{E}W_i = 0$. Now, if in addition we assume $|x_i| < d < \infty \ \forall i$, $P(|\varepsilon_i| \leq m) = 1 \ \forall i$, where $m \in (0,\infty)$, and $\mathrm{var}(\varepsilon_i) = \sigma^2 < \infty$ (which is then implied) we have that

$$\mathrm{var}\left(W_i\right) = \sigma^2 x_i^2 \leq \sigma^2 d^2 < \infty \forall i.$$

Also $P(|x_i\varepsilon_i| \leq dm) = P(|w_i| \leq dm) = 1 \ \forall i$, and $\sum_{i=1}^{n} \mathrm{var}\left(W_i\right) = \sigma^2 \sum_{i=1}^{n} x_i^2 \to \infty$ as $n \to \infty$ (under our previous assumption regarding consistency). Then by Theorem 5.32,

$$\left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \frac{\left(\hat{\beta}_n - \beta\right)}{\sigma} \xrightarrow{d} N\left(0,1\right), \text{ and by Definition 5.2,}$$

$$\hat{\beta}_n \overset{a}{\sim} N\left(\beta, \sigma^2\left(\sum_{i=1}^{n} x_i^2\right)^{-1}\right).$$

d) $\hat{\beta} = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \dfrac{92,017}{897,235} = .1026.$

9. a) To be consistent with the general linear model framework the model can be rearranged using a logarithmic transformation, i.e.,

$Y_t^* = \ln Y_t = \ln \beta_1 + \beta_2 \ln p_t + \beta_3 \ln i_t + \varepsilon_t, t = 1,\ldots,17$,

The statistical model is $\{f(\mathbf{y}^*;\Theta), \Theta \in \Omega\}$ where $\mathbf{y}^* = \left\{y_1^*,\ldots,y_{17}^*\right\}$;

$$\Theta = \begin{bmatrix} \beta^* \\ -- \\ \sigma \end{bmatrix} = \begin{bmatrix} \ln \beta_1 \\ \beta_2 \\ \beta_3 \\ --- \\ \sigma \end{bmatrix} \text{ and } \Omega = \left\{\left(\beta^*,\sigma\right): \beta^* \in R^3, \sigma > 0\right\}; \text{ and}$$

$$f\left(\mathbf{y}^*;\Theta\right) = \prod_{t=1}^{17} \frac{1}{\left(2\pi\right)^{1/2} \sigma} \exp\left(-\frac{1}{2}\left(\frac{y_t^* - \mathbf{x}_t \beta^*}{\sigma}\right)^2\right),$$

with $\mathbf{x}_t = [1 \;\; \ln p_t \;\; \ln i_t]$.

Since the $\mathrm{E}\,Y_t^*$ changes with $t$=1,...,17, $Y^*$ cannot be interpreted as a random sample from a population distribution but, rather, as a random sample generated by a composite experiment.

b) The approach used in the proof of Theorem 8.11, p. 465, can be followed to show that $\mathbf{x}'\mathbf{Y}^*$ and $\mathbf{Y}^{*\prime}\mathbf{Y}^*$ represent a set of minimal, complete sufficient statistics.

c)
- The BLUE estimator of $\beta^*$ is $\hat{\beta} = \left(\mathbf{x}'\mathbf{x}\right)^{-1} \mathbf{x}'\mathbf{Y}^*$.
- The estimate (calculated in GAUSS) is
$$b = \begin{bmatrix} 3.1635 \\ -.8288 \\ 1.1431 \end{bmatrix}.$$

d)
- Because $\hat{\beta} = \left(\mathbf{x}'\mathbf{x}\right)^{-1} \mathbf{x}'\mathbf{Y}^*$ is unbiased under the classical assumptions of the GLM and $\hat{\beta}$ is a function of the complete sufficient statistics, by the Lehmann-Scheffé Completeness Theorem, $\hat{\beta}$ is the MVUE for $\beta^*$.
- See part c) for the estimate of $\beta^*$.

e)
- By Theorem 7.12, the MVUE of $\left(\beta_2 + \beta_3\right)$ is $\left(\hat{\beta}_2 + \hat{\beta}_3\right)$.

- The estimate of the degree of homogeneity of the demand function in terms of relative prices and real income is
  $(b_2 + b_3) = .3143$.

f)
- $\hat{\beta} \sim \mathrm{N}(\beta^*, \sigma^2(\mathbf{x}'\mathbf{x})^{-1})$, which is the probability distribution of the MVUE for ($\ln\beta_1, \beta_2, \beta_3$).

- By Theorem 4.9, $\hat{\beta}_2 + \hat{\beta}_3 = \ell'\hat{\beta} \sim N\left(\beta_2 + \beta_3, \sigma^2 \ell'(\mathbf{x}'\mathbf{x})^{-1}\ell\right)$ where $\ell = [0\ 1\ 1]'$, defines the probability distribution for the MVUE of $\beta_2 + \beta_3$.

g)
- Because the classical assumptions of the GLM hold, if $(\mathbf{x}'\mathbf{x})^{-1} \rightarrow [\mathbf{0}]$ as $n \rightarrow \infty$ then $\hat{\beta}$ is a consistent estimator of $\beta^*$ (see Theorem 8.4).

- By Theorem 8.7, sufficient conditions for asymptotic normality are:

  i) the classical assumptions of the GLM hold
  ii) $\varepsilon_i$'s are iid
  iii) $\boldsymbol{P}(|\varepsilon_i| < m) = 1$ for $m < \infty$ and $\forall i$
  iv) $|x_{ij}| < \xi < \infty \ \ \forall i, j$
  v) $\lim_{n\to\infty} n^{-1}\mathbf{x}'\mathbf{x} = \mathbf{Q}$ a positive definite symmetric matrix.

h)
- $$\hat{S}^2 = \frac{\left(\mathbf{Y}^* - \mathbf{x}\hat{\beta}\right)'\left(\mathbf{Y}^* - \mathbf{x}\hat{\beta}\right)}{n-k} = \frac{\mathbf{Y}'^*\mathbf{Y}^* - \mathbf{Y}'^*\mathbf{x}\left(\mathbf{x}'\mathbf{x}\right)^{-1}\mathbf{x}'\mathbf{Y}^*}{n-k}.$$

  Since $\mathrm{E}\left(\hat{S}^2\right) = \sigma^2$ and $\hat{S}^2$ is a function of complete sufficient statistics

  $\Rightarrow$ $\hat{S}^2$ is the MVUE for $\sigma^2$.

- The MVUE estimate is $\hat{s}^2 = .002366$.

i) Under the classical assumptions of the GLM and if the $\varepsilon_i$'s are iid; then, by Theorem 8.5, $\hat{S}^2$ is a consistent estimator of $\sigma^2$.

j) The appropriate probability estimate is

$$P\left(|b_2 - \beta_2| \leq 1.5\right) = P\left(\left|\frac{b_2 - \beta_2}{\sqrt{\mathrm{var}\left(\hat{\beta}_2\right)}}\right| \leq \frac{.15}{\sqrt{.003173}}\right) = P(|z| \leq 2.6628) = .99225 \quad \text{where } Z \sim N(0,1)$$

and $\hat{\beta}_2 \sim N\left(\beta_2, \ell'\sigma^2\left(\mathbf{x}'\mathbf{x}\right)^{-1}\ell\right)$ with $\ell = (0\ 1\ 0)'$.

k) Since, $\text{plim}\,\hat{\text{E}}Y_t = \text{plim}\left[e^{\hat{\beta}_1^*} p_t^{\hat{\beta}_2} i_t^{\hat{\beta}_3}\right] = e^{\text{plim}\hat{\beta}_1^*} p_t^{\text{plim}\hat{\beta}_2} i_t^{\text{plim}\hat{\beta}_3} = e^{\beta_1^*} p_t^{\beta_2} i_t^{\beta_3} = \beta_1 p_t^{\beta_2} i_t^{\beta_3}$

$\Rightarrow$ $\hat{\text{E}}Y_t$ is a consistent estimator of E$Y_t$ for given values of $p_t$, $i_t$.

l) Because $\ln Y_t \sim N(\mathbf{x}_t\beta, \sigma^2)$; then by problem 9a) in Chapter 6,

$\Rightarrow Y_t \sim$ log-normal with mean $e^{(x_t\beta+\sigma^2/2)}$. So,

$$\hat{\text{E}}Y_t = \text{E}\left[e^{\hat{\beta}_1^*} p_t^{\hat{\beta}_2} i_t^{\hat{\beta}_3}\right] = \text{E}\left[e^{\left(\hat{\beta}_1^* + \hat{\beta}_2 \ln p_t + \hat{\beta}_3 \ln i_t\right)}\right] = \text{E}\left[e^{\left(x_t\hat{\beta}\right)}\right]$$
$$= \text{E}e^{x_t\beta + x_t(\mathbf{x'x})^{-1}\mathbf{x}'\varepsilon} = e^{x_t\beta}\text{E}e^{x_t(\mathbf{x'x})^{-1}\mathbf{x}'\varepsilon} = e^{x_t\beta}e^{\left(\sigma^2 x_t(\mathbf{x'x})^{-1}\mathbf{x}_t'\right)/2}.$$

where the last term follows from the fact that

$\mathbf{x}_t(\mathbf{x}'\mathbf{x})^{-1}\varepsilon \sim N(0, \sigma^2\, \mathbf{x}_t(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}_t')$, and the approach of Problem 9a), Chapter 6 can be applied again.

Since $\hat{\text{E}}Y_t$ is biased, it is neither BLUE nor MVUE.

11. a) $Y_i^* = \ln Y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ is in GLM form. The original parameters $\beta_1$ and $\beta_2$ will be estimated by least squares.

b) If $x_i$'s are not all equal (they of course should not be in this case), then the **x** matrix consisting of a column of 1's and a column of $x_i$ values has full column rank. $\varepsilon_i$'s have the distribution of a Gamma($a$, $b$) random variable that has its mean, $ab$, subtracted from it. Thus $\text{E}\varepsilon_i = 0$, $\forall i$. Since the $\varepsilon_i$'s are iid, $\text{E}\varepsilon\varepsilon' = (ab^2)\mathbf{I}$. Thus, the conditions for the Gauss-Markov Theorem are met, and so the least squares estimator is unbiased and BLUE. The estimator will be asymptotically unbiased if conditions hold such that $n^{1/2}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left([\mathbf{0}], \sigma^2\mathbf{Q}^{-1}\right)$. See Table 8.1 and c) below.

c) Yes, if $|x_{ij}| < \xi < \infty \forall i, j$, in which case the conditions in Table 8.1 for consistency and asymptotic normality would hold.

d) Yes, since the variance of the $\varepsilon_i$'s is $ab^2$, and all of the conditions in Table 8.1 hold for $\hat{S}^2$ to be unbiased and consistent for the variance of the $\varepsilon_i$'s.

13. a) The statistical model can be defined as $\{f(\mathbf{x};p), p \in \Omega\}$ where

$$f(\mathbf{x};p) = \prod_{i=1}^{n} p(1-p)^{x_i - 1} I_{\{1,2,3,\ldots\}}(x_i) = p^n (1-p)^{\sum_{i=1}^n x_i - n} \prod_{i=1}^{n} I_{\{1,2,3,\ldots\}}(x_i),$$

and $\Omega = (0,1)$. This assumes that the geometric distribution characterizes the probability distribution for the number of trials required to obtain the first light.

b) Let $\ln L(p;\mathbf{x}) = n\ln(p) + \left(\sum_{i=1}^{n} x_i - n\right)\ln(1-p) + \ln\prod_{i=1}^{n} I_{\{1,2,3\ldots\}}(x_i)$, then

$$\frac{d\ln L(p;\mathbf{x})}{dp} = \frac{n}{p} + \left(\frac{\sum_{i=1}^{n} x_i - n}{1-p}\right)(-1) = 0$$

$$\Rightarrow \hat{p} = n\Bigg/\left(\sum_{i=1}^{n} X_i\right) = \left(\bar{X}_n\right)^{-1}.$$

(Note: $\frac{d^2\ln L(p;\mathbf{x})}{dp^2} = -\frac{n}{p^2} - \frac{\left(\sum_{i=1}^{n} x_i - n\right)}{(1-p)^2} < 0 \Rightarrow L(p;\mathbf{x})\text{max at }\hat{p}.$)

Since $\mu=1/p$; then, by the ML invariance principle, $t_1(\mathbf{X}) = 1/\hat{p} = \bar{X}_n$ is the ML estimator of $\mu$.

- Since

$$\mathrm{E}\left[t_1(\mathbf{X})\right] = \mathrm{E}\left[\sum_{i=1}^{n} X_i/n\right] = \frac{1}{p}.$$

$\Rightarrow t_1(\mathbf{X})$ is an unbiased estimator of $\mu$. Since $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for $f(\mathbf{x};p)$ / $L(p;\mathbf{x})$, and the MLE is unique and unbiased, then the MLE is MVUE (Theorem 8.14).

- plim $(t_1(\mathbf{X}))$ = plim $\bar{X}_n = \mu = \frac{1}{p}$ (by Khinchin's WLLN)

$\Rightarrow t_1(\mathbf{X})$ is a consistent estimator of $\mu = 1/p$.

- From LLCLT (Theorem 5.30),

$$\frac{n^{1/2}\left(\bar{X}_n - \mu\right)}{\sigma} \xrightarrow{\text{d}} N(0,1)$$

$$\Rightarrow \bar{X}_n \overset{a}{\sim} N\left(\mu, \sigma^2/n\right) = N\left(\frac{1}{p}, \frac{1-p}{np^2}\right),$$

$$\Rightarrow t_1(\mathbf{X}) = \bar{X}_n \overset{a}{\sim} N\left(\frac{1}{p}, \frac{1-p}{np^2}\right).$$

- From Definition 7.22, since $t_1(\mathbf{X})$ is consistent, and the CRLB for estimating $q(p) = p^{-1}$ is

$$\left[\frac{dq(p)}{dp}\right]^2\left[-\mathrm{E}\left[\frac{d^2\ln f(\mathbf{x};p)}{dp^2}\right]\right]^{-1}=\left[\frac{-1}{p^2}\right]^2\left[\frac{n}{p^2(1-p)}\right]^{-1}=\frac{1-p}{np^2}.$$

$\Rightarrow t_1(\mathbf{X})$ is an asymptotically efficient estimator.

c) • The MLE for $p$ is $\hat{p}=\left(\bar{X}_n\right)^{-1}$.

• $\operatorname{plim}\hat{p}=\left(\operatorname{plim}\bar{X}_n\right)^{-1}=(\mu)^{-1}=\left(\frac{1}{p}\right)^{-1}=p$ by Khinchin,s WLLN. Hence $\hat{p}$ is a consistent estimator.

• By LLCLT, $n^{1/2}\left(\bar{X}_n-\frac{1}{p}\right)\xrightarrow{d}N\left(0,\frac{1-p}{p^2}\right)$. Let $g\left(\bar{X}_n\right)=\left(\bar{X}_n\right)^{-1}$ and note that $\frac{dg(\mu)}{d\bar{x}_n}=-\left(\frac{1}{p}\right)^{-2}=-p^2$ is continuous and is nonzero for $p\neq 0$. Also, $\sigma^2\neq 0$ if $p\neq 1$. For $p\epsilon(0,1)$, Theorem 5.39 implies that $\hat{p}=\left(\bar{X}_n\right)^{-1}\overset{a}{\sim}N\left(p,n^{-1}p^2(1-p)\right)$.

• By Definition 7.22, since $\hat{p}$ is consistent, and the CRLB for estimating $q(p)=p$ is given by

$$\left[\frac{dq(p)}{dp}\right]^2\left[-\mathrm{E}\left[\frac{d^2\ln f(x;p)}{dp^2}\right]\right]^{-1}=(1)\left[\frac{n}{p^2(1-p)}\right]^{-1}=\frac{p^2(1-p)}{n}$$

$\Rightarrow$ $\hat{p}$ is an asymptotically efficient estimator.

• Although $\hat{p}$ is an asymptotically efficient and a consistent estimator of $p$, it is not MVUE. We show this below.

Consider $\hat{\hat{p}}=(n-1)\Big/\left(\left(\sum_{i=1}^{n}X_i\right)-1\right)$ as an estimator for $p$. Notice that

$$Y=\sum_{i=1}^{r}X_i\sim f(y;r,p)=\binom{y-1}{r-1}p^r(1-p)^{y-r}I_{\{r,r+1,r+2,\ldots\}}(y),$$

which is the negative binomial PDF. Thus,

$$\mathrm{E}\left[\hat{\hat{p}}\right]=\mathrm{E}\left[\frac{r-1}{Y-1}\right]=\sum_{i=r}^{\infty}\frac{r-1}{i-1}\frac{(i-1)!}{(r-1)!(i-r)!}p^r(1-p)^{i-r}=\sum_{i=r}^{\infty}\frac{(i-2!)}{(r-2)!(i-r)!}p^r(1-p)^{i-r}$$

$$=\sum_{i=r}^{\infty}\frac{(i-2)!}{(s-1)!(i-s-1)!}p^{s+1}(1-p)^{i-s-1}\qquad(\text{substitute } s=r-1),$$

$$=p\sum_{i=s}^{\infty}\frac{(i-1)!}{(s-1)!(i-s)!}p^s(1-p)^{i-s}=p\qquad(\text{note the sume of the negative binomial PDF}),$$

$$\Rightarrow \mathrm{E}\left[\hat{\hat{p}}\right]=p.$$

Since $\hat{\hat{p}}$ is unbiased and a function of the complete sufficient statistic, $\sum_{i=1}^{n}X_i$ ; then, by the Lehmann-Scheffé Completeness Theorem, $\hat{\hat{p}}$ is the MVUE for $p$. By uniqueness, the MLE of $p$ is not the MVUE.

d) For the MLE

- $t_1(x)=\sum_{i=1}^{n}x_i/n=\dfrac{10{,}118}{10{,}000}=1.0118$ is the estimated expected number of trials needed for the first light.
- $\hat{p}=(\bar{x}_n)^{-1}=.9883$ is the estimated probability that the lighter lights on any given trial.

For the MVUE

- $t_2(x)=\left(\sum_{i=1}^{n}x_i-1\right)\Big/(n-1)=1.0118$ is the estimated expected number of trials needed for the first light.
- $\hat{\hat{p}}=(n-1)\Big/\left(\sum_{i=1}^{n}x_i-1\right)=.9883$ is the estimated probability that the lighter lights an any given trial.

Since $t_1(X)$ and $\hat{p}$ are asymptotically efficient they are "approximately" MVUE. Hence, for a large sample size the MLE and the MVUE estimates will be close to each other; here they are identical to the fourth decimal place.

15. a) • The joint density of the random sample $X_1,\ldots,X_n$ is

$$f(\mathbf{x};\Theta)=\prod_{i=1}^{n}\frac{1}{\Theta}I_{(0,\Theta)}(x_i)=\frac{1}{\Theta^n}\prod_{i=1}^{n}I_{(0,\Theta)}(x_i)=L(\Theta;\mathbf{x}).$$

For the $L(\Theta;\mathbf{x})$ to be a maximum, $1/\Theta$ must be as large as possible *while still maintaining* $\prod_{i=1}^{n} I_{(0,\Theta)}(x_i)=1$. Thus, the MLE for $\Theta$ is $\hat{\Theta}_n(\mathbf{X})=\max\{X_1,...,X_n\}$. By Problem 3 part f) in Chapter 7, $\max\{X_1,...,X_n\}$ is a complete sufficient statistic. Hence, it is a minimal sufficient statistic by Theorem 7.7.

- $\mathrm{E}[X_i]=\Theta/2$, which suggests $t_n(\mathbf{X})=\hat{\Theta}_n/2=\max\{X_1,...,X_n\}/2$, is the MLE for the expected number of minutes, $\mu=\Theta/2$, by the invariance principle.

  Since $t_n(\mathbf{X})$ is an invertible function of a complete sufficient statistic, then $t_n(\mathbf{X})$ is a complete, minimal sufficient statistic (Theorem 7.10).

b) • A ML estimate of $\Theta$ is $\hat{\Theta}_n=\max\{x_1,...,x_n\}=13.8$.

- A ML estimate of $\mu=\Theta/2$ is $t_n(\mathbf{x})=\hat{\Theta}_n/2=6.9$.

c) • Since $E\hat{\Theta}_n=\mathrm{E}\left[\max\{X_1,...,X_n\}\right]=\Theta\left(\frac{n}{n+1}\right)$, $\hat{\Theta}_n$ is biased. However, $\lim_{n\to\infty}\mathrm{E}\hat{\Theta}=\Theta$.
  In addition,

$$\lim_{n\to\infty}\mathrm{var}\left(\hat{\Theta}_n\right)=\lim_{n\to\infty}\mathrm{E}\hat{\Theta}_n^2-\lim_{n\to\infty}\left(\mathrm{E}\hat{\Theta}_n\right)^2=\lim_{n\to\infty}\Theta^2\frac{n}{n+2}-\left[\lim_{n\to\infty}\Theta\frac{n}{n+1}\right]^2=0.$$

Hence, $\hat{\Theta}_n\xrightarrow{\mathrm{m}}\Theta\Rightarrow\hat{\Theta}_n\xrightarrow{\mathrm{p}}\Theta$, and $\hat{\Theta}_n$ is consistent.

- Since $\mathrm{E}\left[t_n(\mathbf{X})\right]=\frac{1}{2}\Theta\left(\frac{n}{n+1}\right)\Rightarrow t_n(\mathbf{X})$ is biased. As above, $\lim_{n\to\infty}\mathrm{E}\left[t_n(\mathbf{X})\right]=\Theta/2$ and

$$\lim_{n\to\infty}\mathrm{var}\left(t_n(\mathbf{X})\right)=\lim_{n\to\infty}\mathrm{E}\left[\left(\frac{\hat{\Theta}_n}{2}\right)^2\right]-\lim_{n\to\infty}\left(\mathrm{E}\left[\frac{\hat{\Theta}_n}{2}\right]\right)^2=\lim_{n\to\infty}\frac{1}{4}\Theta^2\left(\frac{n}{n+2}\right)-\lim_{n\to\infty}\frac{1}{4}\Theta^2\left(\frac{n}{n+1}\right)^2=0.$$

Hence, $t_n(\mathbf{X})\xrightarrow{\mathrm{m}}\Theta/2\Rightarrow t_n(\mathbf{X})\xrightarrow{\mathrm{p}}\Theta/2$, so $t_n$ is consistent. Alternatively, since $\hat{\Theta}_n\xrightarrow{\mathrm{p}}\Theta$, $\hat{\Theta}_n/2\xrightarrow{\mathrm{p}}\Theta/2$ by Theorem 5.5.

d) Both $\hat{\Theta}_n$ and $t_n(\mathbf{X})$ are biased $\Rightarrow$ neither are MVUE.

17. a) • Let $(X_1,...,X_n)$ be a random sample from a normal population distribution. The moment conditions can be specified as

$$\mathrm{E}g\left(X_t,\mu,\sigma^2\right)=\mathrm{E}\begin{bmatrix} X_t-\mu \\ X_t^2-\left(\mu^2+\sigma^2\right)\end{bmatrix}=[\mathbf{0}],$$

where $\mathrm{E}\left[X_t^2\right]=\mathrm{var}\left[X_t\right]+\left[\mathrm{E}X_t\right]^2=\sigma^2+\mu^2$. The sample analogs to the moment conditions are

$$n^{-1}\sum_{t=1}^{n} g\left(x_t,\mu,\sigma^2\right)=\begin{bmatrix} m_1'-\mu \\ m_2'-\left(\mu^2+\sigma^2\right)\end{bmatrix}=[\mathbf{0}].$$

The MOM estimator is defined by solving the preceding vector equation as

$$\begin{bmatrix}\hat{\mu} \\ \hat{\sigma}^2\end{bmatrix}=\begin{bmatrix} M_1' \\ M_2'-\left(M_1'\right)^2\end{bmatrix}.$$

- Since $\begin{bmatrix}\mu \\ \sigma^2\end{bmatrix}$ is a continuous function $\forall\left(\mu_1',\mu_2'\right)\in\Gamma \Rightarrow$ the MOM estimator of $\begin{bmatrix}\mu \\ \sigma^2\end{bmatrix}$ is consistent (see Theorem 8.24, p. 499).

- Let the $(i,j)$ element of matrix $\mathbf{A}$ be $\dfrac{\partial h_i^{-1}\left(\mu_1',\mu_2'\right)}{\partial \mu_j'}$

$$\Rightarrow \mathbf{A}=\begin{bmatrix} 1 & 0 \\ -2\mu_1' & 1\end{bmatrix}$$ which consists of continuous elements and has full rank $\forall\left(\mu_1',\mu_2'\right)$.

Then by Theorem 8.25, p. 500, the MOM estimator is asymptotically normally distributed.

b) Let $(X_1,\ldots,X_n)$ be a random sample from a Beta population distribution. The moment conditions can be specified as

$$\mathrm{E}g\left(X_t;\alpha,\beta\right)=\mathrm{E}\begin{bmatrix} X_t-\mu \\ X_t^2-\left(\mu^2+\sigma^2\right)\end{bmatrix}=\begin{bmatrix} \mu_1'-\alpha/\left(\alpha+\beta\right) \\ \mu_2'-\left(\left(\dfrac{\alpha}{\alpha+\beta}\right)^2+\dfrac{\alpha\beta}{\left(\alpha+\beta+1\right)\left(\alpha+\beta\right)^2}\right)\end{bmatrix}=[\mathbf{0}].$$

The sample analogs of the moment conditions are

$$n^{-1}\sum_{t=1}^{n} g\left(x_t;\alpha,\beta\right)=\begin{bmatrix} m_1' - \dfrac{\alpha}{\alpha+\beta} \\ m_2' - \left(\left(\dfrac{\alpha}{\alpha+\beta}\right)^2 + \dfrac{\alpha\beta}{\left(\alpha+\beta+1\right)\left(\alpha+\beta\right)^2}\right) \end{bmatrix} = [\mathbf{0}].$$

Solving for $\hat{\alpha}$:

$$\Rightarrow \hat{\alpha} = \frac{m_1'\hat{\beta}}{1-m_1'}.$$

Solving for $\hat{\beta}$:

$$\Rightarrow m_2' = \left(m_1'\right)^2 + \left(m_1'\right)^2\left(\frac{\hat{\beta}}{\hat{\alpha}^2+\hat{\alpha}\hat{\beta}+\hat{\alpha}}\right)$$

$$\Rightarrow \left(\hat{\alpha}^2+\hat{\alpha}\hat{\beta}+\hat{\alpha}\right)\left(m_2'-\left(m_1'\right)^2\right) = \hat{\beta}\left(m_1'\right)^2$$

$$\Rightarrow \left(\left(\frac{m_1'}{1-m_1'}\right)^2\hat{\beta} + \left(\frac{m_1'}{1-m_1'}\right)\hat{\beta} + \frac{m_1'}{1-m_1'}\right)\left(m_2'-\left(m_1'\right)^2\right) = \left(m_1'\right)^2 \quad \left(\text{sub for } \hat{\alpha} \text{ and divide by } \hat{\beta}\right)$$

$$\Rightarrow \hat{\beta} = m_1' - 1 + \frac{\left(1-m_1'\right)^2\left(m_1'\right)^2}{m_1'\left(m_2'-\left(m_1'\right)^2\right)} = \frac{\left(m_1'-1\right)\left(m_2'-m_1'\right)}{m_2'-\left(m_1'\right)^2}$$

$$\Rightarrow \hat{\alpha} = \frac{-m_1'\left(m_2'-m_1'\right)}{m_2'-\left(m_1'\right)^2}.$$

Since $\begin{bmatrix}\alpha \\ \beta\end{bmatrix}$ is a continuous function $\forall\left(\mu_1',\mu_2'\right)\in\Gamma$, the MOM estimator is consistent (see Theorem 8.24). The **A** matrix of Theorem 8.25 is given by

$$\mathbf{A} = \begin{bmatrix} -\left(\mu_2'\right)^2 + 2\mu_1'\mu_2' - \left(\mu_1'\right)^2\mu_2' & \left(\mu_1'\right)^2\left(\mu_1'-1\right) \\ -4\mu_1'\mu_2' + \left(\mu_1'\right)^2\mu_2' + \left(\mu_2'\right)^2 + \mu_2'\left(\mu_2'+1\right) & -\left(\mu_1'\right)^3 + 2\left(\mu_1'\right)^2 - \mu_1' \end{bmatrix} \div \left(\mu_2' - \left(\mu_1'\right)^2\right)^2.$$

The elements of **A** are continuous functions of $\mu$ and the matrix is nonsingular, so asymptotic normality holds.

c) • Let $(X_1,\ldots,X_n)$ be a random sample from a geometric population distribution. The moment condition can be specified as

$$\mathrm{E}g\left(X_t,p\right) = \mathrm{E}\left[X_t - \frac{1}{p}\right] = \mu_1' - \frac{1}{p} = 0..$$

The sample moment condition is

$$n^{-1}\sum_{t=1}^{n} g\left(x_t, p\right) = \left[ m_1' - \frac{1}{p} \right] = 0.$$

The MOM estimator is defined as

$$\hat{p} = \left( \sum_{i=1}^{n} X_i / n \right)^{-1} = \left( \bar{X}_n \right)^{-1}.$$

By Khinchin's WLLN (Theorem 5.19),

$\bar{X}_n \xrightarrow{p} \mu$ or $\bar{X}_n \xrightarrow{p} \frac{1}{p}$, so $\hat{p} \xrightarrow{p} p$ since $\left(\bar{X}_n\right)^{-1} \xrightarrow{p} \left(\frac{1}{p}\right)^{-1} = p$ by Theorem 5.5, p. 242.

By LLCLT (Theorem 5.30),

$$n^{1/2}\left( \bar{X}_n - \frac{1}{p} \right) \xrightarrow{d} N\left( 0, \frac{1-p}{p^2} \right) \text{ or } \bar{X}_n \overset{a}{\sim} N\left( \frac{1}{p}, \frac{1-p}{np^2} \right).$$

Then since $\frac{d\hat{p}}{d\bar{x}_n} = -\left(\bar{x}_n\right)^{-2}$ is continuous in a neighborhood of $\bar{x}_n = \frac{1}{p}$, it follows from Theorem 5.39 (p. 274) that

$$\hat{p} \overset{a}{\sim} N\left( p, \frac{p^2\left(1-p\right)}{n} \right).$$

d) Let $(X_1,...,X_n)$ be a random sample from a continuous uniform population distribution. The moment conditions are

$$\mathrm{E}g\left(X_t, a, b\right) = \mathrm{E}\begin{bmatrix} X_t - \left( \frac{\left(a+b\right)}{2} \right) \\ X_t^2 - \left[ \frac{\left(b-a\right)^2}{12} + \left( \frac{a+b}{2} \right)^2 \right] \end{bmatrix} = [\mathbf{0}].$$

The sample moment conditions are

$$n^{-1}\sum_{t=1}^{n} g\left(x_t, a, b\right) = \begin{bmatrix} m_1' - \left( \frac{a+b}{2} \right) \\ m_2' - \left[ \frac{\left(b-a\right)^2}{12} + \left( \frac{a+b}{2} \right)^2 \right] \end{bmatrix}.$$

The MOM estimators are derived from

$$m_1' - \left(\frac{a+b}{2}\right) = 0,$$
$$m_2' - 1/3\left(b^2 + ab + a^2\right) = 0,$$

(solve first equation ) $\Rightarrow a = 2m_1' - b$,

(substitute into second equation) $\Rightarrow b^2 - 2m_1'b + \left(4\left(m_1'\right)^2 - 3m_2'\right) = 0.$

Use the quadratic formula (p. 165) to solve for *b*, yielding

$$b = m_1' \pm \sqrt{3}\left(m_2' - \left(m_1'\right)^2\right)^{1/2}.$$

Since b>a, only the root based on + is relevant, yielding

$$\hat{b} = m_1' + \sqrt{3}\left(m_2' - \left(m_1'\right)^2\right)^{1/2}$$
$$\hat{a} = m_1' - \sqrt{3}\left(m_2' - \left(m_1'\right)^2\right)^{1/2}.$$

- Since *a* and *b* are continuous functions of $\left(\mu_1', \mu_2'\right), \forall\left(\mu_1', \mu_2'\right) \in \Gamma,$ the MOM estimator is consistent (Theorem 8.24).
- The **A** matrix of Theorem 8.25 is given by

$$\mathbf{A} = \begin{bmatrix} 1 + (.75)^{1/2}\left(\mu_2' - \left(\mu_1'\right)^2\right)^{-1/2}\left(2\mu_1'\right) & -(.75)^{1/2}\left(\mu_2' - \left(\mu_1'\right)^2\right)^{-1/2} \\ 1 - (.75)^{1/2}\left(\mu_2' - \left(\mu_1'\right)^2\right)^{-1/2}\left(2\mu_1'\right) & (.75)^{1/2}\left(\mu_2' - \left(\mu_1'\right)^2\right)^{-1/2} \end{bmatrix}.$$

The elements of **A** are continuous functions of $\left(\mu_1', \mu_2'\right)\forall\left(\mu_1', \mu_2'\right) \in \Gamma.$ The matrix has full rank, since its determinant is given by $2(.75)^{1/2}\left(\mu_2' - \left(\mu_1'\right)^2\right)^{-1/2} > 0.$ Thus, the MOM is asymptotically normally distributed.

**Chapter 9 – Student Answer Key – Odd Numbers Only**
**Elements of Hypothesis-Testing Theory**

1. a) $H_0 = \left\{ \dfrac{\binom{k}{x}\binom{m-k}{n-x}}{\binom{m}{n}}, k \le 2 \right\}$ and $C_r = \{x: x=2\}$.

$P(\text{type I}) = P(x \in C_r; k) = P(x=2; k) = \dfrac{\binom{k}{2}\binom{20-k}{2-2}}{\binom{20}{2}}$ for $k$=0, 1, or 2,

$P(x=2|k=0) = 0$, $P(x=2|k=1) = 0$ , $P(x=2|k=2) = .005263$.

b) $P(\text{type II}) = P(x \notin C_r; k) = P(x \le 1; k) = P(x=0;k) + P(x=1;k)$ for $k$=3,4,...,20.

| k | P(type II) | k | P(type II) |
|---|---|---|---|
| 3 | .9842 | 12 | .6526 |
| 4 | .9684 | 13 | .5895 |
| 5 | .9474 | 14 | .5211 |
| 6 | .9211 | 15 | .4474 |
| 7 | .8895 | 16 | .3684 |
| 8 | .8526 | 17 | .2842 |
| 9 | .8105 | 18 | .1947 |
| 10 | .7632 | 19 | .1000 |
| 11 | .7105 | 20 | 0 |

c) $\pi(k) = P(x \in C_r; k)$,
$= P(x=2; k)$ for $k \in \{0,1,...,20\}$.

| k | π(k) | K | π(k) |
|---|---|---|---|
| 0 | 0 | 11 | .2895 |
| 1 | 0 | 12 | .3474 |
| 2 | .0053 | 13 | .4105 |
| 3 | .0158 | 14 | .4789 |
| 4 | .0316 | 15 | .5526 |
| 5 | .0526 | 16 | .6316 |
| 6 | .0789 | 17 | .7158 |
| 7 | .1105 | 18 | .8053 |
| 8 | .1474 | 19 | .9000 |
| 9 | .1895 | 20 | 1.000 |
| 10 | .2368 | | |

The graph would plot $\pi(k)$ versus $k$.

There is a very low probability of rejecting the shipment of sets given there are $k \leq 2$ defective sets. The probability that a department store accepts a shipment given there are $k>2$ defective sets is high for $k=3$ and slowly decreases thereafter.

The test has little power for detecting a situation where more than 10% of the units are defective unless the large majority of the units are defective. Thus, the test does not protect the department store from grossly unacceptable shipments very well, while the manufacturer has a low probability of having a shipment returned unless a large percentage of the shipment is defective.

3. a) Both hypotheses are composite, since neither $H_0$ nor $H_a$ identify a unique Gamma distribution.

b) The null hypothesis is simple, denoting the unique geometric distribution having $p$=.01. The alternative hypothesis is composite, and identifies an uncountably infinite number of geometric distributions.

c) Assuming $H_0$: $\beta=[\mathbf{0}]$ and $H_a$: $\beta\neq[\mathbf{0}]$, then both hypotheses are composite since neither hypothesis identifies a unique normal distribution (in the case of $H_0$, although $\beta=[\mathbf{0}]$ specifies a unique value for $\beta$, $\sigma^2$ can be any positive real number).

d) Both hypotheses are simple, since each hypothesis identifies a unique Poisson density.

5. a)

$$H_0 = \left\{ \prod_{i=1}^{50} \frac{1}{B(a,1)} x_i^{a-1} I_{(0,1)}(x_i),\, \mu \le .75 \right\} = \left\{ \exp\left[ -n \ln B(a,1) + (a-1)\sum_{i=1}^{50} \ln(x_i) \right] \prod_{i=1}^{50} I_{(0,1)}(x_i),\, a \le 3 \right\},$$

Note that $\mu \le .75 \leftrightarrow a \le 3$.

The joint density is in the exponential class of densities, where $g(\mathbf{x}) = \sum_{i=1}^{n} \ln(x_i)$ and $c(a) = a-1$. Here $dc(a)/da = 1>0 \Rightarrow$ the joint density has a monotone likelihood ratio in the statistic $g(\mathbf{x})$ (Theorem 9.5).

By Theorem 9.6 and its corollaries, with $H_0$: $a \le 3$ and $H_a$: $a>3$, $C_r = \left\{ x : \sum_{i=1}^{n} \ln(x_i) \ge c \right\}$, for the choice of c such that $P\left( \sum_{i=1}^{n} \ln(x_i) \ge c; a = 3 \right) = \alpha$, is a UMP level .05 test. Because $X_i \sim$ Beta$(a,1) = \dfrac{1}{B(a,1)} x_i^{a-1} I_{(0,1)}(x_i)$, then by a change of variables the distribution of $Y_i = -\ln X_i$ can be derived as (note that $B(a,1)^{-1} = a$; see p. 200)

$$h(y_i) = a\left(e^{-y_i}\right)^{a-1} |-e^{-y_i}|\, I_{(0,\infty)}(y_i) = a e^{-a y_i} I_{(0,\infty)}(y_i),$$

which is an exponential distribution or Gamma(1, 1/$a$). By Theorem 4.2, $-\sum_{i=1}^{n} \ln X_i \sim \text{Gamma}(n, 1/a)$. Solving for $c$ that defines $C_r = \left\{ x : -\sum_{i=1}^{n} \ln(x_i) \le -c \right\}$ with level .05:

$\displaystyle\int_0^{-c} \frac{a^n}{\Gamma(n)} y^{n-1} e^{-ay}\, dy = .05$, where $n$=50 and $a$=3. Using NLSYS in GAUSS, we find $c = -12.9883$.

Thus,

$$C_r = \left\{ x : \sum_{i=1}^{50} \ln x_i \ge c \right\} = \left\{ x : e^{\sum_{i=1}^{50} \ln x_i} \ge e^c \right\} = \left\{ x : \left( \prod_{i=1}^{50} x_i \right)^{1/50} \ge e^{c/50} \right\} = \left\{ x : \left( \prod_{i=1}^{50} x_i \right)^{1/50} \ge e^{-.2598} \right\}.$$

Since, $\bar{x}_g = \left( \prod_{i=1}^{50} x_i \right)^{1/50} = .84 > e^{-.2598} = .7712 \Rightarrow$ reject $H_0$.

b) The power function for the test is

$$\pi(a) = P(x \in C_r; a) = \int_0^{12.9883} \frac{a^{50}}{\Gamma(50)} y^{49} e^{-ay}\, dy\,.$$

Some selected values of $\pi(a)$ are

| a | π(a) | μ=a/(a+1) |
|---|---|---|
| 2.0 | ≈0 | .6667 |
| 2.5 | .0026 | .7143 |
| 3.0 | .0500 | .75 |
| 3.5 | .2692 | .7778 |
| 4.0 | .6254 | .80 |
| 5.0 | .9761 | .8333 |

¶

A potential investor in a restaurant would likely prefer the given hypothesis, which protects against accepting $\mu$>.75 when in fact $\mu$≤.75. On the other hand, from the perspective of a chamber of commerce trying to attract new businesses, a null hypothesis stating that $\mu$ is greater than or equal to .75 may be preferred.

c) The $P$ value for the test is

$$P \text{ value } = \arg\min_{\alpha}\left[-\sum_{i=1}^{n} \ln x_i \;\; \le -c(\alpha)\right]$$

$$= \min_{\alpha}\left[\sup_{a \in H_0} P(x \in C_r(\alpha);\; a) \text{ s.t. } -\ln(.84)^{50} \le -c(\alpha)\right]$$

$$= \int_0^{-\ln(.84)^{50}} \frac{(3)^{50}}{\Gamma(50)} y^{49} e^{-3y} dy$$

≈0, which is strong evidence against $H_0$.

7. In this case, $t(X) = \sum_{i=1}^{100} X_i \sim \text{Binomial } (10, .4)$ under $H_0$, and outcomes of $t(\mathbf{X})$ can be used to characterize the critical region. Note that

| c | P(t(x) = c) |
|---|---|
| 0 | .006047 |
| 1 | .04 |
| 2 | .121 |
| 3 | .215 |
| 4 | .251 |
| 5 | .201 |
| 6 | .111 |
| 7 | .042 |
| 8 | .011 |
| 9 | .001573 |
| 10 | .000105 |

In order to define $a$ ≤.10 size critical region, it is clear that none of the values 2,3,4,5,6 can be in the critical region, or else the size will exceed .10. The probability of the remaining points {0,1,7,8,9,10} equals .101, so that one or more of the points need to be part of the critical region. Also, the critical region cannot be one-sided, or else the test will surely be biased for some $p \in [0,1]$.

One trivial unbiased test is the case where $C_r = \emptyset$. So $P(x \in C_r; .4) = 0 \leq P(x \in C_r; p)\ \forall p \neq .4$. But this test has no power to detect $H_a$.

For any of the remaining possible critical regions to define an unbiased test, they must satisfy the following first order condition for a minimum of the power function at $p$=.4

$$\frac{d\pi(p)}{dp} = \sum_{x \in C_r} \binom{10}{x} (.4)^{x-1} (.6)^{9-x} (x-4) = 0 .$$

None of the possible choices of $C_r$ consisting of subsets of {0,1,7,8,9,10} with probability $\leq$.10 (under $H_0$) satisfy this condition. Thus, except for the useless trivial test presented above, there is no unbiased test with size $\leq$.10.

9. a) $$H_0 = \left\{ \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} I_{\{0,1,2,\ldots\}}(x_i), \lambda \geq 3 \right\},$$

$$= \left\{ \exp\left[ -n\lambda - \sum_{i=1}^{n} \ln(x_i!) + \ln(\lambda) \sum_{i=1}^{n} x_i \right] \prod_{i=1}^{n} I_{\{0,1,2,\ldots\}}(x_i),\ \lambda \geq 3 \right\}.$$

The joint density is in the exponential class of densities with $g(\mathbf{x}) = \sum_{i=1}^{n} x_i$ and $c(\lambda) = \ln(\lambda)$. Since $dc(\lambda)/d\lambda = 1/\lambda > 0$, then by Theorem 9.5 the joint density has a monotone likelihood ratio in the statistic $g(\mathbf{x})$.

By Theorem 9.6 and its corollaries, if $H_0$: $\lambda \geq 3$ and $H_a$: $\lambda < 3$, $C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \leq c \right\}$ for a choice of $c$ such that $P\left( \sum_{i=1}^{n} x_i \leq c; \lambda = 3 \right) = \alpha$ is the type of critical region needed.

Because the $X_i$'s are distributed iid Poisson with mean $\lambda$, implying $\sum_{i=1}^{n} X_i$ is distributed Poisson with mean $n\lambda$ (straightforwardly shown using the MGF approach; see answer 4a), select $c$ such that

$$P\left( \sum_{i=1}^{12} x_i \leq c; \lambda = 3 \right) = \sum_{i=0}^{c} \frac{e^{12 \cdot 3} (12 \cdot 3)^i}{i!} = \sum_{i=0}^{c} \frac{e^{-36} (36)^i}{i!},$$

will be as close to .10 as possible without exceeding .10.

Some specific possibilities for the size of the test are given below.

| **c** | $P\left(\sum_{i=1}^{12} x_i \le c;\ \lambda = 3\right)$ |
|---|---|
| 26 | .0513 |
| 27 | .0736 |
| 28 | .1023 |
| 29 | .1379 |
| 30 | .1806 |

Hence, $C_r = \left\{ \mathbf{x}: \sum_{i=1}^{12} x_i \le 27 \right\}$.

b) Since $C_r = \left\{\mathbf{x}: \sum_{i=1}^{12} x_i \le 27\right\} = \{\mathbf{x}: \bar{x} \le 2.25\}$ and $\bar{x} = 2 \Rightarrow$ reject $H_0$. Thus, the safety program was effective in reducing the average accidents per week at the .0736 level of significance.

11. $C_r = \left\{ \mathbf{x} : \frac{ns^2}{\sigma_0^2} \le \chi^2_{n-1;1-\alpha} \right\}$.. This can be proven by reversing inequalities in the derivation presented in Example 9.26.

13. a) If the null hypothesis is $H_0$: $\theta \geq 7.5$, then a test of level $\alpha$ would be protecting against rejecting components that meet the minimum mean operating life requirements, and sufficiently small observations of mean life will be required to obtain high enough power to reject a false $H_0$. This may protect sales of the company, but may not be wise in terms of public safety and long run reputation of the company, which would be protected by setting up an $\alpha$-level test for the null hypothesis $H_0$: $\theta \leq 7.5$.

b) We protect the company's sales in setting up the test below, i.e., we test $H_0$: $\theta \geq .7.5$, but see a) above.

$$H_0 = \left\{ \prod_{i=1}^{n} \frac{1}{\theta} e^{-x_i/\theta} I_{(0,\infty)}(x_i), \theta \ge 7.5 \right\}$$
$$= \left\{ \exp\left[ -n \ln\theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i \right] \prod_{i=1}^{n} I_{(0,\infty)}(x_i), \theta \ge 7.5 \right\},$$

which is in the exponential class of densities with $c(\theta) = -1/\theta,\ g(\mathbf{x}) = \sum_{i=1}^{n} x_i$. Since $dc(\theta)/d\theta > 1/\theta^2 > 0$, then by Theorem 9.5 the joint density has a monotone likelihood ratio in the statistic $g(\mathbf{x})$.

By Theorem 9.6 and its corollaries, the UMP critical region of size $\alpha$ for testing $H_0$: $\theta \geq 7.5$ versus $H_a$: $\theta < 7.5$ is represented by

$C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \le c \right\}$ for $c$ chosen so that $P(\mathbf{x} \in C_r;\ \theta{=}7.5) = \alpha$.

To find $c$, recall that $\sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \theta)$,

$$\int_0^c \frac{1}{\theta^n \Gamma(n)} x^{n-1} e^{-x/\theta} dx = \alpha,$$

where $n$=300 in this application.

Using the NLSYS application package in GAUSS with $\alpha$= .01, yields $c$=1958.85. Here, $\alpha$ was chosen as .01 to keep type I error low and require strong evidence against $H_0$ before denying the company the contract. (But recall the discussion above regarding public safety!).

c) Given that $\alpha$=.01, the critical region is defined as

$$C_r = \left\{ \mathbf{x} : \sum_{i=1}^{300} X_i \le 1958.85 \right\} = \left\{ \mathbf{x} : \bar{x} \le 6.53 \right\}.$$

Since $\bar{x} = \sum_{i=1}^{300} x_i / 300 = 7.8 \notin C_r,$ fail to reject $H_0$.

Based on the random sample, the company's electronic component does not violate the minimum mean operating life of the component. Hence, it should be awarded the contract.

15. If $\sigma^2$ is known to equal $\sigma_*^2$, then the joint density of the random sample can be represented as

$$f(\mathbf{x};\mu) = \frac{1}{(2\pi)^{n/2}\sigma_*^n} \exp\left( -\frac{1}{2\sigma_*^2} \left( \sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2 \right) \right).$$

This is in exponential class form (see p. 220) with $\mathrm{c}(\mu) = \mu / \sigma_*^2$ and $g(\mathbf{x}) = \sum_{i=1}^{n} x_i$. Since $dc(\mu)/d\mu > 0$, the joint density has a monotone likelihood ratio in the statistic $\sum_{i=1}^{n} x_i$. Then Theorem 9.6 and its corollaries can be used to answer parts a) and b).

a) $C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \ge c \right\}$, where $C_r$ is chosen so that $P(\mathbf{x} \in C_r;\ \mu{=}\mu_0) = \alpha$.

b) $C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \le c \right\}$, , where $C_r$ is chosen so that $P(\mathbf{x} \in C_r;\ \mu{=}\mu_0) = \alpha$.

Both of the preceding critical regions could have also used set defining conditions in terms of values of the sample mean.

c) To test the two sided hypothesis of part c), Theorem 9.9 (p. 571) could be used, leading to

$$C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \le c \text{ or } \sum_{i=1}^{n} x_i \ge c_2 \right\},$$

where $c_1$ and $c_2$ are chosen such that $P(\mathbf{x} \epsilon C_r; \mu{=}\mu_0) = \alpha$ and $d\pi_{C_R}(\mu_0)/d\mu = 0.$ This $C_r$ could be defined in terms of values of $\bar{x}$ as well.

17. a) $H_0 = \left\{ \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} I_{\{0,1\}}(x_i), p = .032 \right\}.$

The joint density is in the exponential class of densities with $c(p) = \ln\left(p/(1-p)\right)$, $g(\mathbf{x}) = \sum_{i=1}^{n} x_i$, $d(p) = n\ln(1-p)$, $z(\mathbf{x}) = 0$, and $A = \times_{i=1}^{n} \{0,1\}$.

By Theorem 9.9, since $dc(p)/dp = [p(1-p)]^{-1} > 0$, a UMPU level-$\alpha$ test of the hypothesis $H_0$: $p$=.032 versus $H_a$: $p \ne .032$ has critical region

$$C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \le c_1 \text{ or } \sum_{i=1}^{n} x_i \ge c_2 \right\},$$

for choices of $c_1$ and $c_2$ satisfying $d\pi_{C_r}(.032)/dp = 0.$

For large enough $n$, an asymptotic normal distribution of the test statistic under $H_0$ is derived from

$$Z = \frac{\sum_{i=1}^{n} X_i - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\bar{X} - p_0}{\left[n^{-1}p_0(1-p_0)\right]^{1/2}} \xrightarrow{d} N(0,1)$$ defines the critical region

$$C_r^* = \left\{ \mathbf{x} : z \le -z_{\alpha/2} \text{ or } z \ge z_{\alpha/2} \right\}.$$

b) For $\alpha$=.05, $C_r^* = \{ \mathbf{x} : z \le -1.96 \text{ or } z \ge 1.96 \}$.

Since

$$z_1 = \frac{.034 - .032}{\left[\dfrac{.032(1-.032)}{100}\right]^{1/2}} = -.1136 \Rightarrow$$ fail to reject $H_0 : p = .032.$ The process is under control.

c) In this case

$$z_1 = \frac{.026 - .032}{\left[\frac{.032(1-.032)}{100}\right]^{1/2}} = -.3409 \Rightarrow \text{ fail to reject } H_0 : p = .032.$$

There is insufficient evidence that the quality has increased based on the outcome of the test. The new proportion is not statistically different from .032. A one-sided test would lead to the same conclusion.

19. a) $$H_0 = \left\{ \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} I_{\{0,1\}}(x_i), \mu \in [2,4] \right\}.$$

The joint density is in the exponential class of densities with $c(p) = \ln(p/(1-p))$ and $g(\mathbf{x}) = \sum_{i=1}^{n} x_i$. By Theorem 9.9, since $dc(p)/dp = [p(1-p)]^{-1} > 0$, a UMPU level-$\alpha$ test of the hypothesis $H_0$: $p \in [.02, .04]$ versus $H_a$: $p \notin [.02, .04]$ is given by the critical region

$$C_r = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \le c_1 \text{ or } \sum_{i=1}^{n} x_i \ge c_2 \right\},$$

for choices of $c_1$ and $c_2$ that satisfy $\pi_{C_r}(.02) = \pi_{C_r}(.04) = \alpha$.

For large enough $n$, an asymptotic normal distribution of the test statistic is derived from

$$\frac{\sum_{i=1}^{n} X_i - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\bar{X} - p_0}{\left[n^{-1} p_0 (1-p_0)^{1/2}\right]^{1/2}} \xrightarrow{\text{d}} N(0,1),$$

and defines the critical region

$$C_r^* = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \le c_1 \text{ or } \sum_{i=1}^{n} x_i \ge c_2 \right\},$$

for choices of $c_1$ and $c_2$ that satisfy $1 - \pi_{C_r}(.02) = 1 - \pi_{C_r}(.04) = 1 - \alpha$.

For $\alpha$=.05, and with $Z = \sum_{i=1}^{n} X_i \overset{a}{\sim} N\left(z; np_0, np_0(1-p_0)\right), n = 400,$ and $p_0 = .02$ or $.04$,

$$\int_{c_1}^{c_2} N(z; 8, 7.84)\, dz = \int_{c_1}^{c_2} N(z; 16, 15.36)\, dz = 1 - .05,$$
$$\Rightarrow C_r^* = \left\{ \mathbf{x} : \sum_{i=1}^{400} x_i \le 3.394 \text{ or } \sum_{i=1}^{400} x_i \ge 22.471 \right\} = \left\{ \mathbf{x} : \bar{x} \le .00849 \text{ or } \bar{x} \ge .05618 \right\}.$$

The values for $c_1$ and $c_2$ were found using the NLSYS procedure in GAUSS.

b) From the random sample of $400, \bar{x} = .07 \in C_r^*$.

$\Rightarrow$reject $H_0$ and conclude that the vehicle should not be classified in the average risk class.

c) Since .07 is higher than the upper bound of the average risk class, one could conclude that this sedan is in the high risk class. A UMPU level-$\alpha$ test of a null hypothesis designed to test whether or not the vehicles' frequency of claims are consistent with the high risk category could be implemented, but will not contradict this conclusion.

**Chapter 10 – Student Answer Key – Odd Numbers Only**
**Hypothesis-Testing Methods**

1. a) The likelihood function is

$$L(\mu,\sigma;\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}.$$

Let $\theta = \left[\mu,\sigma^2\right]'$. The GLR test is

$$\lambda(x) = \frac{L(\mu_0 = 16.03, \sigma_0 = .01; \mathbf{x})}{\sup\limits_{\theta\in H_0\cup H_1} L(\mu,\sigma;\mathbf{x})}.$$

To find the maximum value of the denominator, we take the first order conditions

$$\frac{\partial \ln L(\mu,\sigma;\mathbf{x})}{\partial\mu} = \frac{\partial}{\partial\mu}\left[\frac{-n}{2}\ln\left(2\pi\sigma^2\right) + \left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right)\right)\right]$$

$$= \left(-2\sum_{i=1}^{n} x_i + 2n\mu\right)\left(-\frac{1}{2\sigma^2}\right) = 0 \Rightarrow \hat{\mu} = \overline{x} = 15.8943,$$

$$\frac{\partial \ln L(\mu,\sigma;\mathbf{x})}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{2\pi\sigma^2}2\pi + \frac{1}{2\sigma^4}\left(\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right) = 0$$

$$= -\frac{n}{2\sigma^2} + \frac{n}{2\sigma^4}\left(\sum_{i=1}^{n} x_i^2 / n - 2\mu\overline{x} + \mu^2\right) = 0$$

$$\Rightarrow \hat{\sigma}^2 = \sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)^2 / n = (.1015)^2$$

$$\Rightarrow \lambda(\mathbf{x}) = \frac{L(16.03,\ .01;\ x)}{L(15.8943,\ .1015;\ x)} \approx 0$$

$$\Rightarrow -2\ln\left(\lambda(\mathbf{X})\right) > 10^6.$$

From Theorem 10.5 $-2\ln\left(\lambda(\mathbf{x})\right) \sim \chi_1^2$ and $\chi^2_{1;.10} = 2.706$. Then for $\alpha = .10 \Rightarrow$ reject $H_0$. Thus, the filling process is not under control at mean fill rate of $\mu = 16.03$ with standard deviation $\sigma = .01$.

b) Let $\theta = \left[\mu,\sigma^2\right]'$. Since $\hat{\theta}_r = [16.03,\ .0001]'$ and

$$\frac{\partial^2 \ln L(\theta;\mathbf{x})}{\partial\theta\partial\theta'} = n\left[\begin{array}{c|c} -1/\sigma^2 & -(\bar{x}-\mu)/\sigma^4 \\ \hline -(\bar{x}-\mu)/\sigma^4 & \dfrac{1}{2\sigma^4} - \dfrac{1}{\sigma^6}\left(\sum_{i=1}^{n} x_i^2/n - 2\mu\bar{x} + \mu^2\right) \end{array}\right],$$

then

$$w = \frac{\partial \ln L(\hat{\theta}_r;\mathbf{x})'}{\partial\theta}\left[\frac{-\partial^2 \ln L(\hat{\theta}_r;\mathbf{x})}{\partial\theta\,\partial\theta'}\right]^{-1}\frac{\partial \ln L(\hat{\theta}_r;\mathbf{x})}{\partial\theta}$$

$$= \begin{bmatrix}-54,280 & 57,275,000\end{bmatrix}\left(-\begin{bmatrix}-40,000 & 5.43\times10^8 \\ 5.43\times10^8 & -1.1475\times10^{12}\end{bmatrix}\right)^{-1}\begin{bmatrix}-54,280 \\ 57,275,000\end{bmatrix} = 8029.6,$$

$\Rightarrow$ reject $H_0$ since for $\alpha = .10$ the critical value is $\chi^2_{2;.10} = 4.605$. Yes, the GLR and LM tests are in agreement.

c) Using the GLR approach, we test $H_0: \mu = 16.03$ and $H_0: \sigma = .01$ individually, using size .05 tests. For $H_0: \mu = 16.03$, we have

$$\lambda(x) = \sup_{\theta\in H_0} L(\mu,\sigma;\mathbf{x}) \Big/ \sup_{\theta\in H_0\cup H_2} L(\mu,\sigma;\mathbf{x}).$$

To determine the supremum of the likelihood in the numerator

$$\frac{\partial L(\mu_0,\sigma;\mathbf{x})}{\partial\sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu_0)^2 = 0$$

$$\Rightarrow \hat{\sigma}_0^2 = \sum_{i=1}^{n}(x_i-\mu_0)^2/n$$

$$\Rightarrow \lambda(x) = L(\mu_0,\hat{\sigma}_0,\mathbf{x})/L(\hat{\mu},\hat{\sigma};\mathbf{x}),$$

where $\hat{\mu},\hat{\sigma}$ are defined in part a)
$\Rightarrow \lambda(\mathbf{x}) \approx 0 \Rightarrow -2\ln(\lambda(\mathbf{x})) > 10^6 > \chi^2_{1;.05} = 3.841 \Rightarrow$ reject $H_0$. For $H_0: \sigma = .01$, we have

$$\lambda(\mathbf{x}) = \sup_{\theta\in H_0} L(\mu,\sigma;\mathbf{x}) \Big/ \sup_{\theta\in H_0\cup H_2} L(\mu,\sigma;\mathbf{x}).$$

To determine the supremum of the likelihood in the numerator

$$\frac{\partial \ln L(\mu,\sigma_0;\mathbf{x})}{\partial\mu} = n(\bar{x}-\mu)/\sigma_0^2 = 0 \Rightarrow \hat{\mu} = \bar{x}.$$

Thus, $\lambda(x) = \frac{L(\hat{\mu},\sigma_0;\mathbf{x})}{L(\hat{\mu},\hat{\sigma};\mathbf{x})} \approx 0 \Rightarrow -2\ln(\lambda(x)) > 10^6 > \chi^2_{1;.05} = 3.841 \Rightarrow$ reject $H_0$.

The Bonferroni inequality provides an upper bound to type **I** error if the above two hypothesis are considered collectively. That is,

$$\mathbf{P}(\text{type } \mathbf{I} \text{ error}) \leq \sum_{i=1}^{2} \alpha_i = .10,$$

which implies that jointly the conclusion that $H_0 : \mu = 16.03$ and $\sigma = .01$ is protected against a type one error at level .10.

d) To form a Wald test of $H_0 : \mu = 16.03, \sigma = .01$ versus $H_a :$ not $H_0$, we examine the equivalent hypothesis $H_0 : \mu = 16.03, \sigma^2 = .0001$ versus $H_a :$ not $H_0$. From Theorem 10.9, and the discussion following proof of Theorem 8.18,

$$n\hat{\Sigma}_n = n\left( -\frac{\partial^2 \ln L(\hat{\theta};\mathbf{x})}{\partial\theta\, \partial\theta'} \right)^{-1} = \left[ \begin{array}{c|c} 1/\hat{\sigma}^2 & (\bar{x}-\hat{\mu})/\hat{\sigma}^4 \\ \hline (\bar{x}-\hat{\mu})/\hat{\sigma}^4 & \frac{-1}{2\hat{\sigma}^4} + \frac{1}{\hat{\sigma}^6}\sum_{i=1}^{n}(x_j - \hat{\mu})^2 / n \end{array} \right]^{-1},$$

is a consistent estimator of $\Sigma$.

Letting $\hat{\theta} = \begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{bmatrix} = \begin{bmatrix} 15.8943 \\ .0103 \end{bmatrix}$, $R(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\theta$, $r = \begin{bmatrix} \mu_0 \\ \sigma_0^2 \end{bmatrix} = \begin{bmatrix} 16.03 \\ .0001 \end{bmatrix}$, and $\frac{\partial \mathbf{R}(\theta)}{\partial\theta} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then the Wald statistic is calculated as

$$w = \begin{pmatrix} \hat{\mu} - \mu_0 \\ \hat{\sigma}^2 - \sigma_0^2 \end{pmatrix}' \left( -\frac{\partial^2 \ln L(\hat{\theta};\mathbf{x})}{\partial\theta\partial\theta'} \right) \begin{pmatrix} \hat{\mu} - \mu_0 \\ \hat{\sigma}^2 - \sigma_0^2 \end{pmatrix} = 91.1137.$$

Since $w > \chi^2_{2;.10} = 4.605$, reject $H_0$.

3. a) $H_0 : \mu_1 \leq \mu_2$ versus $H_a :$ not $H_0$.

Suppose $X_{1i} \sim N(\mu_1, \sigma^2)$ and $X_{2i} \sim N(\mu_2, \sigma^2)$. Let $\theta = (\mu_1, \mu_2, \sigma^2)$ and define the GLR as $\lambda(\mathbf{x}) = \sup_{\theta \in H_0} L(\theta, \mathbf{x}) / \sup_{\theta \in H_0 \cup H_a} L(\theta, \mathbf{x})$,

where $\ln L(\theta;\mathbf{x}) = \sum_{i=1}^{2}\left[ -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_{ij} - \mu_i)^2 \right].$

To determine the supremum of the denominator

$$\frac{\partial \ln L}{\partial \mu_i} = -\frac{1}{2\sigma^2}\sum_{j=1}^{n} 2\left(x_{ij}-\mu_i\right)(-1) = 0 \Rightarrow \hat{\mu}_i = \sum_{j=1}^{n} x_{ij}/n = \overline{x}_i \quad \text{for } i = 1,2 \text{ and,}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = \frac{-n}{\sigma^2} + \frac{1}{2\sigma^4}\left(\sum_{j=1}^{n}\left(x_{1j}-\mu_1\right)^2 + \sum_{j=1}^{n}\left(x_{2j}-\mu_2\right)^2\right) = 0$$

$$\Rightarrow \hat{\sigma}^2 = \left(\sum_{j=1}^{n}\left(x_{1j}-\hat{\mu}_1\right)^2 + \sum_{j=1}^{n}\left(x_{2j}-\hat{\mu}_2\right)^2\right)/2n.$$

To determine the supremum of the numerator, solve

$$\max\left\{\ln L\left(\mu_1,\mu_2,\sigma^2;\mathbf{x}\right) + \gamma\left(\mu_1-\mu_2\right)\right\},$$

which has first-order conditions (using the Kuhn-Tucker approach)

$$\frac{\partial \ln L}{\partial \mu_i} = -\frac{1}{2\sigma^2}\sum_{j=1}^{n} 2\left(x_{ij}-\mu_i\right)(-1) + (-1)^{i+1}\gamma \le 0,\ \mu_i\left(\frac{\partial \ln L}{\partial \mu_i}\right) = 0 \ \text{ for } i = 1,2,$$

$$\frac{\partial \ln L}{\partial \sigma^2} = \frac{-n}{\sigma^2} + \frac{1}{2\sigma^4}\left(\sum_{j=1}^{n}\left(x_{1j}-\mu_1\right)^2 + \sum_{j=1}^{n}\left(x_{2j}-\mu_2\right)^2\right) = 0,$$

$$\frac{\partial \ln L}{\partial \gamma} = \mu_1-\mu_2 \le 0,\ \gamma\left(\frac{\partial \ln L}{\partial \gamma}\right) = 0.$$

If $\hat{\mu}_1 < \hat{\mu}_2$, so that the constraint is not binding, then $\gamma = 0$ by complementary slackness

$$\Rightarrow \hat{\mu}_i = \sum_{j=1}^{n} x_{ij}/n \quad \text{for } i = 1,2,$$

with

$$\hat{\sigma}^2 = \left(\sum_{j=1}^{n}\left(x_{1j}-\hat{\mu}_1\right)^2 + \sum_{j=1}^{n}\left(x_{2j}-\hat{\mu}_2\right)^2\right)/2n.$$

If $\hat{\mu}_1 = \hat{\mu}_2 = \hat{\mu}$, so that the constraint is binding, then the FOCs imply that

$$\hat{\mu} = \overline{x}_1 + \frac{\sigma^2}{n}\gamma \text{ and } \hat{\mu} = \overline{x}_2 - \frac{\sigma^2}{n}\gamma \Rightarrow \hat{\mu} = \left(\overline{x}_1+\overline{x}_2\right)/2,$$

with $\hat{\sigma}_0^2 = \left(\sum_{j=1}^{n}\left(x_{1j}-\hat{\mu}\right)^2 + \sum_{j=1}^{n}\left(x_{2j}-\hat{\mu}\right)^2\right)/2n.$ In defining the GLR there are three distinct cases:

i) $\hat{\mu}_1 < \hat{\mu}_2 \Rightarrow \lambda(\mathbf{x}) = 1,$ because the constraint is not binding;

ii) $\hat{\mu}_1 = \hat{\mu}_2 \Rightarrow \lambda(\mathbf{x}) = 1,$ because $\hat{\mu} = \hat{\mu}_1 = \hat{\mu}_2$;

iii) $\hat{\mu}_1 > \hat{\mu}_2 \Rightarrow \lambda(\mathbf{x}) < 1,$ because the constraint is binding.

The GLR is then

$$\lambda(\mathbf{x}) = \begin{cases} L(\hat{\mu}, \hat{\mu}, \hat{\sigma}_0^2; \mathbf{x}) / L(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}^2; \mathbf{x}) & , \text{ if } \hat{\mu}_1 > \hat{\mu}_2 \\ 1 & , \text{ if } \hat{\mu}_1 \le \hat{\mu}_2 \end{cases}$$

The GLR can be rewritten as

$$\lambda(\mathbf{x}) = \begin{cases} (\hat{\sigma}^2 / \sigma_0^2)^{n/2} & , \text{ if } \hat{\mu}_1 > \hat{\mu}_2 \\ 1 & , \text{ if } \hat{\mu}_1 \le \hat{\mu}_2 \end{cases}$$

Substituting the definitions of $\hat{\mu}, \hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}^2,$ and $\hat{\sigma}_0^2,$ and noting that $\hat{\mu} = ½(\hat{\mu}_1 + \hat{\mu}_2)$ yields

$$\lambda(\mathbf{x}) = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-n/2} \left[1 + \frac{n}{2} \frac{(\hat{\mu}_1 - \hat{\mu}_2)^2}{\left(\sum_{j=1}^{n} (x_{1j} - \hat{\mu}_1)^2 + \sum_{j=1}^{n} (x_{2j} - \hat{\mu}_2)^2\right)}\right]^{-n/2}.$$

- Letting
$$t = \frac{(\hat{\mu}_1 - \hat{\mu}_2) / \sqrt{2/n}}{\sqrt{\left(\sum_{j=1}^{n} (x_{1j} - \hat{\mu}_1)^2 + \sum_{j=1}^{n} (x_{2j} - \hat{\mu}_2)^2\right) / (2n-2)}} \Rightarrow \lambda(\mathbf{x}) = \left(1 + t^2 / (2n-2)\right)^{-n/2}.$$

- $\lambda = 1$ when $\hat{\mu}_1 = \hat{\mu}_2$ and thus $t$=0 and is montonically decreasing as $t$ increases in magnitude $\Rightarrow \lambda \le c$ iff $t \ge c^* > 0.$

Under $H_0$, $T$ has a central $t$-distribution with $2n-2$ degrees of freedom when $\mu_1 = \mu_2$. Thus, the critical value $c^*$ can be found from the $t$-table.

Given $\hat{\mu}_1 = 15.8943,$ $\hat{\mu}_2 = 16.02753,$ $\hat{\sigma}_1^2 = .01031,$ and $\hat{\sigma}_2^2 = .009968,$ the value of the test statistic is

$$t = \frac{\dfrac{(15.8943 - 16.02753)}{(2/40)^{1/2}}}{\left[40[.01031 + .009968]/78\right]^{1/2}} = -5.843 < t_{78;.05} = 1.665,$$
$\Rightarrow$ do not reject $H_0$.

b) The derivation of the test is identical to the approach in 3a) upon reversing the roles of $\mu_1$ and $\mu_2$. The test outcome is

$t = 5.843 > t_{78;.05} = 1.665,$
$\Rightarrow$ reject $H_0 : \mu_2 \leq \mu_1$.

c) In order to define the GLR in this case, first note that the denominator of the ratio is the same as in part 3a). The numerator can be derived following the approach in 3a) except the equality constraint $\mu_1 = \mu_2 = \mu$ replaces the inequality constraint, leading to the pooled mean estimate $\hat{\mu}$ as in 3a), as well as the same estimate of $\hat{\sigma}_0^2$. The GLR is defined as

$$\lambda(\mathbf{x}) = \begin{cases} \left(\sigma^2 / \hat{\sigma}_0^2\right)^{n/2} & \text{if} \quad \hat{\mu}_1 \neq \hat{\mu}_2, \\ 1 & \text{if} \quad \hat{\mu}_1 = \hat{\mu}_2. \end{cases}$$

Using the properties of the likelihood ratio discussed in 3a), it follows that

$\lambda(\mathbf{x}) \leq c$ iff $t \leq c_\ell$ or $t \geq c_u$,

where $t$ is the $t$-statistic defined in 3a). Here, $T$ has a $t$-distribution with $2n-2$ degrees of freedom under $H_0$. The critical region for a size $\alpha$ test could be defined as

$$C_r^T = \left\{ t : t \leq -t_{2n-2,\alpha/2} \text{ and } t \geq t_{2n-2,\alpha/2} \right\}.$$

In this case the test statistic is $t = -5.843$ with critical values $t_{78,\,.025} = 1.991$ so that $t \in C_r$
$\Rightarrow$ reject $H_0 : \mu_1 = \mu_2$.

d) In part a), the power function for $H_0 : \mu_1 \leq \mu_2$ versus $H_a :$ not $H_0$ is given by
$\pi(\delta) = P(t \geq 1.665 \mid \delta)$,

where $\delta = {}^{\mu_1 - \mu_2}\!/\!_{\sqrt{2\sigma^2/n}}$ is the noncentrality parameter. Selected values of $\pi(\delta)$ are given below

| δ | π(δ) |
|---|---|
| -1 | .0042 |
| -1/2 | .0162 |
| 0 | .05 |
| ½ | .1252 |
| 1 | .2567 |
| 2 | .6332 |
| 3 | .9081 |

In part b), the power function for $H_0 : \mu_2 \leq \mu_1$ versus $H_a :$ not $H_0$ is given by
$\pi(\delta) = P(t \geq 1.665 \mid \delta)$,

where $\delta = {}^{\mu_2-\mu_1}\!\big/\!\sqrt{2\sigma^2/n}$ is the noncentrality parameter. Selected power values are given in the preceding table.

In part c), the power function for $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 \neq \mu_2$ is given by

$$\pi(\delta) = P(t \leq -1.991 \mid \delta) + P(t \geq 1.991 \mid \delta),$$

where $\delta$ is defined as in part a) above. Selected values of $\pi(\delta)$ are given in the following table.

| $\delta$ | $\pi(\delta)$ |
|---|---|
| 5.5, −5.5 | .9997 |
| 4, −4 | .9767 |
| 2, −2 | .5062 |
| 1, −1 | .1671 |
| ½, −½ | .0784 |
| 0 | .05 |

For each of the above power functions, the type I error is bounded by .05. Protection against type II error increases as the difference $(\mu_1 - \mu_2) \in H_a$ increased in magnitude. In each case, the difference in mean values must be a multiple of the standard deviation of the difference in sample means for the rejection probability to be appreciable.

5. Examine the case where $H_0 : \sigma^2 \leq d$ and $H_a : \sigma^2 > d$ (the other case can be dealt with analogously). Assume the conditions stated at the end of Section 10.6 ensure that $n^{1/2}\left(\hat{S}^2 - \sigma^2\right) \xrightarrow{d} N\left(0, \mu_4' - \sigma^4\right)$. If $\sigma^2 = d$, then $Z_{(n)} = n^{1/2}\left(\hat{S}^2 - d\right)/\hat{\xi}^{1/2} \xrightarrow{d} N(0,1)$ by Slutsky's theorem (recall the definition of $\hat{\xi}$ at the bottom of p. 639). Under these conditions it follows from the asymptotic distribution of $Z$ that $P\left(z_{(n)} \geq z_\alpha\right) = \alpha$. If $\sigma^2 = \tau < d$, then it follows from the preceding limiting distribution result that $n^{1/2}\left(\hat{S}^2 - \tau\right)/\hat{\xi}^{1/2} \xrightarrow{d} N(0,1)$, so that $Z_{(n)} + n^{1/2}(d-\tau)/\hat{\xi}^{1/2} \xrightarrow{d} N(0,1)$. Then

$P\left(z_{(n)} \geq z_\alpha\right) = P\left(n^{1/2}\left(\hat{s}^2 - \tau\right)/\hat{\xi}^{1/2} \geq z_\alpha + n^{1/2}(d-\tau)/\hat{\xi}^{1/2}\right) =$

$P\left(\left(\hat{s}^2 - \tau\right)/\hat{\xi}^{1/2} \geq n^{-1/2} z_\alpha + (d-\tau)/\hat{\xi}^{1/2}\right) \to 0$ as $n \to \infty$ because

$\left(\hat{S}^2 - \tau\right)/\hat{\xi}^{1/2} \xrightarrow{p} 0/\left(\mu_4' - \sigma^4\right) = 0$, $n^{-1/2} z_\alpha \to 0$, and $(d-\tau)/\hat{\xi}^{1/2} \xrightarrow{p} (d-\tau)/\left(\mu_4' - \sigma^4\right) > 0$.

Thus, the test will have size $\alpha$ asymptotically.

Using an argument similar to the above, the test is consistent since if $\sigma^2 = \tau > d$,

$P\left(z_{(n)} \geq z_\alpha\right) = P\left(\left(\hat{s}^2 - \tau\right)/\hat{\xi}^{1/2} \geq n^{-1/2} z_\alpha + (d-\tau)/\hat{\xi}^{1/2}\right) \to 1$ as $n \to \infty$

Since $\left(\hat{S}^2-\tau\right)/\hat{\xi}^{1/2}\xrightarrow{p}0$ and $\left(d-\tau\right)/\hat{\xi}^{1/2}\xrightarrow{p}\left(d-\tau\right)/\left(\mu_4'-\tau^4\right)<0.$

7. a) The parameter space is given by $\Omega=\left\{\left(\theta_1,\theta_2\right)\mid\theta_1>0,\theta_2>0\right\}$. The likelihood function can be written as

$$L\left(\theta_1,\theta_2;\mathbf{x}\right)=\prod_{i=1}^{n}\frac{1}{\theta_1}e^{-x_{li}/\theta_1}I_{(0,\infty)}\left(x_{li}\right)\prod_{i=1}^{n}\frac{1}{\Theta^2}e^{-x_{2i}/\theta_2}I_{(0,\infty)}\left(x_{2i}\right).$$

The first order conditions for unconstrained maximization yield (for denominator of GLR)

$$\frac{\partial\ln L}{\partial\theta_1}=\frac{\partial}{\partial\theta_1}\left(-n\ln\left(\theta_1\right)-\sum_{i=1}^{n}x_{li}/\theta_1-n\ln\left(\theta_2\right)-\sum_{i=1}^{n}x_{2i}/\theta_2+\ln\left(\prod_{i=1}^{n}I_{(0,\infty)}\left(x_{li}\right)\prod_{i=1}^{n}I_{(0,\infty)}\left(x_{(2i)}\right)\right)\right)$$

$$=-\frac{n}{\theta_1}+\sum_{i=1}^{n}x_{li}/\theta_1^2=0\Rightarrow\hat{\theta}_1=\sum_{i=1}^{n}x_{li}/n=\overline{x}_1,$$

$$\frac{\partial\ln L}{\partial\theta_2}=\frac{-n}{\theta_2}+\sum_{i=1}^{n}x_{2i}/\theta_2^2=0\Rightarrow\hat{\theta}_2=\sum_{i=1}^{n}x_{2i}/n=\overline{x}_2.$$

Under $H_0:\theta_1=\theta_2=\theta$, and after substitution of the constraint the FOC becomes (for numerator of GLR)

$$\frac{\partial\ln L}{\partial\theta}=\frac{-2n}{\theta}+\sum_{j=1}^{n}x_{1j}/\theta^2+\sum_{j=1}^{n}x_{2j}/\theta^2=0\Rightarrow\hat{\theta}=\left(\sum_{j=1}^{n}x_{1j}+\sum_{j=1}^{n}x_{2j}\right)/2n.$$

The GLR is defined as $\lambda\left(\mathbf{x}\right)=1$ if $\hat{\theta}_1=\hat{\theta}_2=\hat{\theta}$. If $\hat{\theta}_1\neq\hat{\theta}_2$, then the GLR is given by

$$\lambda\left(\mathbf{x}\right)=\frac{\left(\frac{1}{\hat{\theta}^n}e^{-\sum_{i=1}^{n}x_{1i}/\hat{\theta}}\right)\left(\frac{1}{\hat{\theta}^n}e^{-\sum_{i=1}^{n}x_{2i}/\hat{\theta}}\right)}{\left(\frac{1}{\hat{\theta}_1^n}e^{-\sum_{i=1}^{n}x_{1i}/\hat{\theta}_1}\right)\left(\frac{1}{\hat{\theta}_2^n}e^{-\sum_{i=1}^{n}x_{2i}/\hat{\theta}_2}\right)}=\frac{\left(\sum_{j=1}^{n}x_{1j}/n\right)^n\left(\sum_{j=1}^{n}x_{2j}/n\right)^n}{\left[\left(\sum_{j=1}^{n}x_{1j}+\sum_{j=1}^{n}x_{2j}\right)/2n\right]^{2n}}\ \left(\text{subbing in for }\hat{\theta},\hat{\theta}_1,\hat{\theta}_2\right),$$

$$=\frac{\left(\overline{x}_1\overline{x}_2\right)^n}{\left[\left(\overline{x}_1+\overline{x}_2\right)/2\right]^{2n}}.$$

To define a size .10 critical region for the test, note that
$\lambda^*\left(\mathbf{x}\right)=\left(\lambda\left(\mathbf{x}\right)\right)^{1/n}\leq c^{1/n}=c^*$ iff $\lambda\left(\mathbf{x}\right)\leq c.$

$$\Rightarrow P\left[\lambda^*(\mathbf{x}) \le c^*\right] = \alpha = P\left[\frac{\bar{x}_1\bar{x}_2}{(\bar{x}_1+\bar{x}_2)^2} \le \frac{c^*}{4}\right] = P\left[\frac{c^*}{4}\left(\bar{x}_1^2 + 2\bar{x}_1\bar{x}_2 + \bar{x}_2^2\right) \ge \bar{x}_1\bar{x}_2\right]$$

$$= P\left[\frac{c^*}{4}\left(\frac{\bar{x}_1}{\bar{x}_2} + 2 + \frac{\bar{x}_2}{\bar{x}_1}\right) \ge 1\right] = P\left[x + \frac{1}{x} \ge \frac{4}{c^*} - 2\right] \text{ where } x = \frac{\bar{x}_1}{\bar{x}_2}$$

$$= P\left[x^2 - dx + 1 \ge 0\right] \text{ where } d = \frac{4}{c^*} - 2.$$

By definition of the GLR, $c \le 1$, so that $c^* \le 1$ and $d \ge 2$. For $d \ge 2$ the roots $r_1, r_2$ of the quadratic form are such that $0 < r_1 \le r_2 < d$.

$$\Rightarrow P\left(x^2 - dx + 1 \ge 0\right) = P(x \le r_1) + P(x \ge r_2) = P(\bar{x}_1 \le r_1\bar{x}_2) + P(\bar{x}_1 \ge r_2\bar{x}_2).$$

Since under $H_0, \bar{x}_1 \sim \text{Gamma}(n, \Theta/n)$ and $\bar{x}_2 \sim \text{Gamma}(n, \Theta/n)$, and $\bar{x}_1, \bar{x}_2$ are independent, the joint density is

$$f(\bar{x}_1, \bar{x}_2) = \left(\frac{1}{(\Theta/n)^n \Gamma(n)} \bar{x}_1^{n-1} e^{-n\bar{x}_1/\theta} I_{(0,\infty)}(\bar{x}_1)\right)\left(\frac{1}{(\Theta/n)^n \Gamma(n)} \bar{x}_2^{n-1} e^{-n\bar{x}_2/\theta} I_{(0,\infty)}(\bar{x}_2)\right).$$

Then

$$P(\bar{x}_1 \le r_1\bar{x}_2) = \int_0^\infty f_2(\bar{x}_2)\left(\int_0^{r_1\bar{x}_2} f_1(\bar{x}_1)\, d\,\bar{x}_1\right) d\,\bar{x}_2 \text{ and } P(\bar{x}_1 \ge r_2\bar{x}_2) = \int_0^\infty f_1(\bar{x}_1)\left(\int_0^{\bar{x}_1/r_2} f_2(\bar{x}_2)\, d\,\bar{x}_2\right) d\,\bar{x}_1.$$

(Notice that both integrals are invariant to the choice of $\theta > 0$. This is because the integrals are improper and are characterized by their limiting behavior at $\infty$).

b) Using the NLYSYS package in GAUSS and the Simpson 1/3 quadrature rule, we solved the two equation system

$$P(\bar{x}_1 \le r_1\bar{x}_2) + P(\bar{x}_1 \ge r_2\bar{x}_2) = \alpha,\ r_1 r_2 = 1,$$

which yielded $r_1 = .7185$ and $r_2 = 1.3917$. Note the second constraint follows from the fact that the two roots of a quadratic equation $ax^2 - dx + c = 0$ have a product equal to $c/a$, which equals 1 in this case. Thus,

$$c^* = \frac{4}{d+2} = \frac{4}{(.7185 + 1.3917) + 2} = .9732,$$

where $d = r_1 + r_2$ from the quadratic formula.

Since $\lambda^* = \left( \dfrac{(24.23)(18.23)}{\left[(24.23+18.23)/2\right]^2} \right) = .98$

$\Rightarrow$ do not reject $H_0$.

Since supplier number one has a lower bid and a chip with an equal operating life, it is reasonable to purchase chips from supplier one.

c) For example, if $H_0 : \theta_1 \le \theta_2$ versus $H_a : \theta_1 > \theta_2$, the GLR would become

$$\lambda(\mathbf{x}) = \begin{cases} \left( \dfrac{\bar{x}_1 - \bar{x}_2}{\left[(\bar{x}_1 + \bar{x}_2)/2\right]^2} \right) & \text{if } \hat{\theta}_1 > \hat{\theta}_2, \\ 1 & \text{if } \hat{\theta}_1 \le \hat{\theta}_2. \end{cases}$$

The GLR test of $H_0 : \theta_1 \le \theta_2$ versus $H_a : \theta_1 > \theta_2$ is a level-$\alpha$ test when $\left(P(\lambda / \mathbf{x}) \le c\right) = \alpha$, and $\lambda(\mathbf{x}) \le c$ iff $\bar{x}_1 / \bar{x}_2 \ge c_*$. In this case the distribution of $\bar{x}_1$ is $\text{Gamma}(n, \theta_1 / n)$ and $\bar{x}_2$ is $\text{Gamma}(n, \theta_2 / n)$.

d) $H_0 : \theta_1 = \theta_2$ versus $H_a : \theta_1 \ne \theta_2$.

The LM test rule is given by $w \begin{bmatrix} \ge \\ < \end{bmatrix} \chi^2_{1;\alpha} \Rightarrow \begin{bmatrix} \text{reject } H_0 \\ \text{accept } H_0 \end{bmatrix}$, where

$$W = \frac{\partial \ln L(\theta_r; \mathbf{x})'}{\partial \theta} \left[ \frac{-\partial^2 \ln L(\hat{\theta}_r; \mathbf{x})}{\partial \theta \, d\theta'} \right]^{-1} \frac{\partial \ln L(\hat{\theta}_r; \mathbf{x})}{\partial \theta}.$$

Let $\hat{\theta}_r = \hat{\theta} = (\bar{x}_1 + \bar{x}_2)/2$. Also note that

$$\frac{\partial^2 \ln L(\hat{\theta}_r; \mathbf{x})}{\partial \theta \, \partial \theta'} = \begin{bmatrix} n / \hat{\theta}_r^2 - 2n\bar{x}_1 / \hat{\theta}_r^3 & 0 \\ 0 & n / \hat{\theta}_r^2 - 2n\bar{x}_2 / \hat{\theta}_r^3 \end{bmatrix}.$$

The LM statistic is thus

$$w = \begin{pmatrix} -n/\theta_r + n\bar{x}_1 / \hat{\theta}_r^2 \\ -n/\theta_r + n\bar{x}_2 / \hat{\theta}_r^2 \end{pmatrix}' \begin{pmatrix} \dfrac{-n}{\hat{\theta}_r^3} + 2n\bar{x}_1 / \hat{\theta}_r^2 & 0 \\ 0 & -\dfrac{n}{\hat{\theta}_r^2} + 2n\bar{x}_2 / \hat{\theta}_r^3 \end{pmatrix}^{-1} \begin{pmatrix} -n/\hat{\theta}_r + n\bar{x}_1 / \hat{\theta}_r^2 \\ -n/\hat{\theta}_r + n\bar{x}_2 / \hat{\theta}_r^2 \end{pmatrix}.$$

Given $\hat{\theta}_r = (\bar{x}_1 + \bar{x}_2)/2 = 21.23 \Rightarrow w = 2.17 < \chi^2_{1;.10} = 2.706 \Rightarrow$ do not reject $H_0$.

e) $H_0 : \theta_1 = \theta_2$ versus $H_a : \theta_1 \neq \theta_2$. The Wald rule is $w \begin{bmatrix} \geq \\ < \end{bmatrix} \chi^2_{1;\alpha} \Rightarrow \begin{bmatrix} \text{reject } H_0 \\ \text{accept } H_0 \end{bmatrix}$, where

$$W = \left[R(\hat{\theta}) - r\right]' \left[ \frac{\partial R(\hat{\theta})}{\partial \theta}' \hat{\Sigma}_n \frac{\partial R(\hat{\theta})}{\partial \theta} \right]^{-1} \left[R(\hat{\theta}) - r\right].$$

Define

$$\frac{\partial^2 \ln L(\hat{\theta};\mathbf{x})}{\partial \theta \partial \theta'} = \begin{pmatrix} n/\hat{\theta}_1^2 - 2n\bar{x}_1/\theta_1^3 & 0 \\ 0 & n/\hat{\theta}_2^2 - 2n\bar{x}_2/\hat{\theta}_2^3 \end{pmatrix}$$

The WALD statistic, with $\hat{\theta}_1 = \bar{x}_1$ and $\hat{\theta}_2 = \bar{x}_2$, is given by

$$w = (\hat{\theta}_1 - \hat{\theta}_2)\left( [1 \quad -1] \begin{pmatrix} n/\hat{\theta}_1^2 & 0 \\ 0 & n/\hat{\theta}_2^2 \end{pmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)^{-1} (\hat{\theta}_1 - \hat{\theta}_2) = 1.96 < \chi^2_{1;.10} = 2.706$$

$\Rightarrow$ do not reject $H_0$.

9. Sorting the values from lowest to highest, it is then evident that a median of the observations is given by $(15.88 + 15.89)/2 = 15.885$. Based on this median value, the number of runs above and below the median in the sequential rowwise observations in Problem 10.1 is equal $w = 19$. In these 19 runs, $n_1 = 20$ observations are above the median and $n_0 = 20$ observations are below. From p. 679, and Example 10.32

$$\mathrm{E}(W) = \frac{2(20)(20)}{40} + 1 \text{ and } \mathrm{var}(W) = \frac{2(20)(20)(2(20)(20) - 20 - 20)}{(20+20)^2(20+20-1)},$$

so that $\mathrm{E}(W) = 21$ and $\mathrm{var}(W) = 9.7436$. Then $z = (w - \mathrm{E}W)/\mathrm{std}(W) = (19-21)/(9.7436)^{1/2} = -.6407$ and the critical region for $Z$ based on its asymptotic $N(0,1)$ distribution for a .05 size test is $C_r = (-\infty, -1.96) \cup (1.96, \infty)$. Since $z \notin C_r$, the hypothesis that the random sample is from some population distribution cannot be rejected.

11. The distinction to be made in this problem is that the observations are independent between pairs, but the observations within a pair should not necessarily be considered independent.

a) The likelihood function for the mean $\mu = \mu_a - \mu_b$ and variance $\sigma^2$ of the population distribution of the differences is

$$L\left(\mu;\sigma^2;\mathbf{d}\right) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(d_i-\mu)^2} = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(d_i-\mu)^2}.$$

where $d_i = x_{ai} - x_{bi}, \sigma^2 = \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}$, and $\mu = \mu_a - \mu_b$. The objective is to test $H_0 : \mu = 0$ versus $H_a :$ not $H_0$. Let $\theta = \left(\mu, \sigma^2\right)'$. The GLR is defined as

$$\lambda(\mathbf{d}) = \sup_{\theta \in H_0} L(\theta;\mathbf{d}) / \sup_{\theta \in H_0 \cup H_a} L(\theta;\mathbf{d}) = \frac{\left(\frac{1}{2\pi\hat{\sigma}_0^2}\right)^{n/2} e^{-\frac{n}{2}}}{\left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{n/2} e^{-\frac{n}{2}}} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2}$$

$$= \frac{1}{1 + \frac{\left(\bar{d}\right)^2}{\sum_{i=1}^{n} d_i^2 / n - \left(\bar{d}\right)^2}} \text{ (substituting for } \hat{\sigma}_0^2, \hat{\sigma}^2, \text{ and rearranging terms),}$$

where $\hat{\sigma}_0^2 = \sum_{i=1}^{n} d_i^2 / n,\ \hat{\sigma}^2 = \sum_{i=1}^{n} \left(d_i - \hat{\mu}\right)^2 / n,\ \hat{\mu} = \bar{x}_a - \bar{x}_b = \bar{d}$.

Under $H_0$ the statistic $t = \bar{d} \big/ \hat{\sigma}/\sqrt{n-1}$ has a $t$-distribution with $(n-1)$ degrees of freedom. The in terms of the $t$-statistic, the test is given by

$$t \begin{bmatrix} \in \\ \notin \end{bmatrix} \left(-\infty, -t_{n-1;\alpha/2}\right] \cup \left[t_{n-1;\alpha/2}, \infty\right) \Rightarrow \begin{bmatrix} \text{reject } H_0 \\ \text{do not reject } H_0 \end{bmatrix}.$$

For $\alpha = .05,\ n = 50,\ \bar{d} = -11.73,$ and $s = 3.89$,

$$t_0 = \frac{-11.73}{3.89/\sqrt{49}} = -21.1080 < -2.011 = -t_{49;.025} \text{ (interpolating in the } t\text{-table)},$$

$\Rightarrow$ reject $H_0$.

Thus, we reject that the meal planning has no effect.

b) Let $\hat{\sigma}_r^2 = \hat{\sigma}_0^2 = \sum_{i=1}^{n} \left(d_i\right)^2 / n$ and $\mu_r = 0$. Then

$$\left.\frac{\partial \ln L}{\partial \theta}\right|_{\theta=\hat{\theta}_r} = \begin{pmatrix} n\,\bar{d}\,/\,\hat{\sigma}_r^2 \\ -\frac{n}{2\hat{\sigma}_r^2} + \frac{1}{2\hat{\sigma}_r^4}\sum_{i=1}^{n}\left(d_i\right)^2 \end{pmatrix},$$

$$\left.\frac{d^2 \ln L}{\partial \theta\, \partial \theta'}\right|_{\theta=\hat{\theta}_r} = n\begin{pmatrix} -1/\hat{\sigma}_r^2 & -\bar{d}\,/\,\sigma_r^4 \\ -\bar{d}\,/\,\hat{\sigma}_r^4 & 1/2\hat{\sigma}_r^4 - 1/\hat{\sigma}_r^6\sum_{i=1}^{n}\left(d_i\right)^2/n \end{pmatrix}.$$

From Theorem 10.7.3 the LM statistic is given by

$$w = \left(\frac{\partial \ln L\left(\hat{\theta}_r;\mathbf{x}\right)}{\partial \theta}\right)'\left(-\frac{\partial^2 \ln L\left(\hat{\theta}_r;\mathbf{x}\right)}{\partial \theta\, \partial \theta'}\right)^{-1}\left(\frac{\partial \ln L\left(\hat{\theta}_r;\mathbf{x}\right)}{\partial \theta}\right) = -56.18.$$

Note that in theory, $w$ should not be negative. This is an example of how, in finite samples, the estimate of the covariance matrix of $\partial \ln L\left(\hat{\Theta}_r;X\right)/\partial\Theta,$ represented by $\frac{\partial^2 \ln l\left(\hat{\theta}_r;\mathbf{x}\right)}{\partial\theta\,\partial\theta'},$ can fail to be positive definite. In this event, the test cannot be used. Regarding additional information, one would seek an alternative method of calculating the covariance matrix. For further discussion and references, see L.G. Godfrey, Misspecification Tests in Econometrics, Cambridge Univ. Press, 1991, p. 80-82.

c) To test if the paired differences, $d_i = x_{ai} - x_{bi},$ are from a normal population once could use a $\chi^2$ goodness-of-fit, Kolmogorov-Smirnov and Lilliefors, or Shapiro-Wilk test. Each of these nonparametric tests utilize the actual observations, which are not available in this problem.

d) An alternative is to specify a Wald test of asymptotic size and use Theorem 10.9. By Theorem 5.37 (Multivariate LLCLT), if

- $\begin{bmatrix} x_a \\ x_b \end{bmatrix}$ are iid with $\mathrm{E}\begin{bmatrix} x_a \\ x_b \end{bmatrix} = \begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}.$
- $\mathrm{cov}\begin{bmatrix} x_a \\ x_b \end{bmatrix} = \Sigma$ where $\Sigma$ is positive definite.

$$\Rightarrow n^{1/2}\left(\bar{\mathbf{x}} - \boldsymbol{\mu}\right) \xrightarrow{\mathrm{d}} N\left(\left[0\right],\Sigma\right).$$

In addition, if $n\Sigma \xrightarrow{p} \Sigma$ and $R\left(\theta\right) = \mu_a - \mu_b,$ then we can define a Wald test of asymptotic size $\alpha$. Given

$$\Sigma = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ba} & \sigma_b^2 \end{pmatrix}, \text{ then by Theorems 6.7 and 6.8 } \Rightarrow n\hat{\Sigma} = n\begin{pmatrix} S_a^2/n & S_{ab}/n \\ S_{ab}/n & S_b^2/n \end{pmatrix} \xrightarrow{\mathrm{p}} \Sigma.$$

Thus,

$$w = [\bar{x}_a - \bar{x}_b]\left([1 -1]\begin{pmatrix} S_a^2/n & S_{ab}/n \\ S_{ab}/n & S_b^2/n \end{pmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)^{-1}[\bar{x}_a - \bar{x}_b] = [\bar{x}_a - \bar{x}_b]\left(\frac{1}{n}\left(S_a^2 - 2S_{ab} + S_b^2\right)\right)^{-1}[\bar{x}_a - \bar{x}_b]$$

$$= [\bar{x}_a - \bar{x}_b]\left(\frac{n}{S^2}\right)[\bar{x}_a - \bar{x}_b] = (-11.73)\left(50/(3.89)^2\right) = 454.639 > 3.84 = \chi^2_{1;\,.05},$$

$\Rightarrow$ reject $H_0$.

13. a) Classifying the sample outcomes as 1's or 0's according to whether the outcome is $\geq$ or $<$ the median (equal to 4) yields $w = 22$ runs $(n_1 = 24, n_0 = 16)$.

Since $n_1$ and $n_0$ are each > 10, we use the normal approximation for $W$. Here

$$z = \frac{w - \mathrm{E}(W)}{\left(\mathrm{var}(W)\right)^{1/2}} = .6013 .$$

For a size .10 runs test, the critical region is $(-\infty, -1.645] \cup [1.645, \infty) \Rightarrow$ do not reject $H_0$.

Thus, we fail to reject that the observations are a random sample from some population distribution.

b) The MLE estimate of $\lambda$, based on the Poisson population distribution assumptions, is $\hat{\lambda} = \bar{x} = 4.45$. We use the $\chi^2$ goodness-of-fit test discussed at the bottom of p. 659 based on the conservative critical value.

For $m = 5$, the number of observations and estimates for $p_i$ for various intervals are given below.

| **i** | **Δi** | $\mathbf{n_i}$ | $\hat{\mathbf{p}}_{\mathbf{i}}$ **(Poisson (4.45))** |
|---|---|---|---|
| 1 | [0,1] | 6 | .0636 |
| 2 | [2,3] | 10 | .2872 |
| 3 | [4,5] | 10 | .3606 |
| 4 | [6,7] | 6 | .2060 |
| 5 | [8,9] | 8 | .0666 |

Since $W = \sum_{i=1}^{5} \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i} = 17.54 > \chi^2_{m-1-k;\,.10} = \chi^2_{3;\,.10} = 6.25$

$\Rightarrow$ reject $H_0$. We reject that the observations are from a Poisson population distribution.

c) For $m = 5$, the number of observations and estimates of $p_i$ for various intervals are given in the table below.

| i | Δi | $n_i$ | $\hat{p}_i$ (Uniform on {0,1,…,9}) |
|---|---|---|---|
| 1 | [0,1] | 6 | .20 |
| 2 | [2,3] | 10 | .20 |
| 3 | [4,5] | 10 | .20 |
| 4 | [6,7] | 6 | .20 |
| 5 | [8,9] | 8 | .20 |

Since $w = 2.00 < \chi^2_{m-1;\,.10} = \chi^2_{4;\,.10} = 7.78$

$\Rightarrow$ do not reject $H_0$. Thus, we cannot reject that the observations were generated from the uniform population distribution $f(x) = \frac{1}{10} I_{\{1,1,2,\ldots,9\}}(x)$.

d) First of all, we recode the outcomes 0,1,…,9 as 1,2,…, 10 to allow a uniform distribution on the outcomes to be represented in the form exhibited in Chapter 4, Section 4.1.1 (there is no loss of generality in doing so). Assuming the observations are from a discrete uniform distribution, then $\mu = (M+1)/2$. Hence, the hypothesis

$H_0 : \mu \geq 8$ versus $H_a : \mu < 8$,

is equivalent to $H_0 : M \geq 15$ versus $H_a : M < 15$. The likelihood function for this problem is

$$L(M;x) = \prod_{i=1`}^{n} \frac{1}{M} I_{\{1,\ldots,M\}}(x_i) = \frac{1}{M^n} \prod_{i=1}^{n} I_{\{1,\ldots,M\}}(x_i).$$

The MLE estimate is $\hat{M} = \max\{x_1,\ldots,x_n\}$. To see this note, that

- for any integer $k < \hat{M}$ then $\prod_{i=1}^{n} I_{\{1,\ldots,\hat{k}\}}(x_i) = 0.$
- for any integer $k > \hat{M}$, then $\frac{1}{k^n} < \frac{1}{\hat{M}^n}.$

To define a size .10 GLR test consider

$$\lambda(\mathbf{x}) = \sup_{M \in H_0} L(M;\mathbf{x}) \Big/ \sup_{M \in H_0 \cup H_a} L(M;\mathbf{x}) = \frac{\frac{1}{\hat{M}_0^n} \prod_{i=1}^{n} I_{\{1,\ldots,M_0\}}(x_i)}{\frac{1}{\hat{M}^n} \prod_{i=1}^{n} I_{\{1,\ldots,\hat{M}\}}(x_i)} = \left(\frac{\hat{M}}{\hat{M}_0}\right)^n.$$

Notice that in fact $\lambda(\mathbf{x})$ is monotonically increasing in the statistic $\hat{M} = \max\{x_1,\ldots,x_n\}$ so that $\lambda(\mathbf{x}) \leq c$ iff $\hat{M} \leq c^*$ $\lambda(\mathbf{x}) = \frac{(10)^{40}}{(15)^{40}} \approx 0$ is the calculated value of the GLR.

To define the critical region, since $\hat{M} = \max\{x_1,...,x_n\}$ is an order statistic, then by the Corollary 6.1 (p. 351)

$$P(\lambda(\mathbf{x}) \le c) = P(\hat{M} \le c^*) = (F(c^*))^{40} = \alpha,$$

where $F(c^*) = \sum_{\substack{z \le c^* \\ z \in \{1,2,...,15\}}} \frac{1}{15} I_{\{1,2,...,15\}}(z)$ (based on $H_0$, with $M_0 = 15$).

Some specific choices of test size are given as follows

| $C^*$ | α |
|---|---|
| 12 | ≈0 |
| 13 | .0033 |
| 14 | .0633 |
| 15 | 1.0 |

In this case a level .10 GLR test occurs when $\alpha = .0633$ and $c^* = 14$, which equates to $c$ = .0633.

Since $\hat{M} = 10 \le 14$, reject $H_0$. In order to generate a .90-level confidence interval for the expected number of daily breakdowns, we need to find all of the values of $M_0$ such that

$$\lambda(\mathbf{x}) = \left(\frac{\hat{M}}{M_0}\right)^{40} > c = .0633,$$

Where $\hat{M} = 10$. This implies that $M_0 < \hat{M} / (.0633)^{1/40} = 10 / .9333 = 10.7143$. Thus, the .90-level confidence estimate for $M$ is [0, 10.7143), which in terms of the expected value of the recoded uniform distribution equates to $\mu \in [0, 5.857)$. In the original units of measurement, the confidence interval for expected breakdowns would be [0, 4.857).

15. a) Classifying the sample outcomes as 1's or 0's according to whether the outcome is > or < 0 yields $w$ = 17 runs $(n_1 = 18, n_0 = 25)$.

Since $n_1$ and $n_0$ are each > 10, we use the normal approximation to $W$. Here

$$z = \frac{w - \mathrm{E}[W]}{(\mathrm{var}(W))^{1/2}} = -1.564 \text{ .}$$

For a size .05 runs test the critical region is $(-\infty, -1.96] \cup [1.96, \infty) \Rightarrow$ do not reject $H_0$.

Since $\mathbf{e} = y - x\hat{\beta}$ and $\operatorname{cov}(\mathbf{e}) = \sigma^2\left(\mathbf{I} - \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\right)$, the $e_i$'s are not iid random variables. However, for large $n$, $e = \left(\mathbf{I} - \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\right)\varepsilon \approx \varepsilon$, and so the runs test may serve as an approximate test for large samples.

b) Test $H_0 : z \sim N(\mu, \sigma^2)$ versus $H_a : z \not\sim N(\mu, \sigma^2)$ using the K-S test with Lilliefors critical values (see p. 674).

For $\hat{\mu} = \overline{\hat{e}} = -.1237$ and $\hat{\sigma}^2 = 1.010$, we use $N(-.1237, 1.010)$ as the distribution of $Z$ under $H_0$. The value of the K-S statistic is $d_n = .0908 < .886/(43)^{1/2} = .1351$.

$\Rightarrow$ do not reject $H_0$ for a size .05 level test.

c) Under normality the expected daily return of the firm is $\mathrm{E}[R_t] = \beta_1 + \beta_2 R_{mt}$. Testing whether or not $\mathrm{E}[R_t]$ is proportional to $R_{mt}$ is equivalent to testing

- $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 \neq 0$.

In terms of a $t$-statistic (with $R$ = [1 0], $r$ = 0)

$$|t| = \left|\frac{Rb - r}{\left(\hat{s}^2 R(\mathbf{x}'\mathbf{x})^{-1} R'\right)^{1/2}}\right| = |2.89| > t_{41;\,.025} = 2.021,$$

$\Rightarrow$ reject $H_0$ when using a size .05 test.

d) $R_t \le R_{mt} \forall R_{mt} \ge 0$ *iff* $\beta_1 \le 0$ and $\beta_2 \le 1$. So test $H_0 : \beta_1 \le 0$ and $\beta_2 \le 1$ versus $H_a$ : not $H_0$. We will use the Bonferroni approach to test this joint hypothesis on the $\beta$-values.

- $H_0 : \beta_1 \le 0$ versus $H_a : \beta_1 > 0$.

  Let $R = [1 \;\; 0]$, $b = [.03 \;\; 1.07]'$, $r = 0$, and then
  $t = (Rb - r)/\left[\hat{s}^2 R(\mathbf{x}'\mathbf{x})^{-1} R'\right]^{1/2} = .03/.01037 = 2.893.$
  Since $t_{41;\,.025} = 2.02$ and the critical region for the test is
  $C_r = [2.02, \infty), t \in C_r \Rightarrow$ reject $H_0$ for $\alpha = .025$.

- $H_0 : \beta_2 \le 1$ versus $H_a : \beta_1 > 1$.

  Let $R = [0 \;\; 1]$, $b = [.03 \;\; 1.07]'$, and $r = 1$, and then
  $t = (Rb - r)/\left[\hat{s}^2 R(\mathbf{x}'\mathbf{x})^{-1} R'\right]^{1/2} = .07/.93323 = .75.$

Since $C_r = [2.02,\infty)$ and $t \notin C_r$ do not reject $H_0$ for $\alpha = .025$.

- The joint hypothesis is thus rejected, so that it can be concluded that $R_t > R_{mt}$ for at least some $R_{mt} > 0$, with size of test $\leq .05$ by Bonferroni's inequality.

17. a) The hypothesis is $H_0 : p_1 = p_2$ versus $H_a : p_1 \neq p_2$. Let

$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \prod_{i=1}^{2} \binom{n_i}{x_i} p_i^{x_i} (1-p_i)^{n_i - x_i} I_{\{,1,\ldots,n_i\}}(x_i)$. In this case

$$\begin{bmatrix} n_1^{1/2} & 0 \\ 0 & n_2^{1/2} \end{bmatrix}\left(\begin{pmatrix} X_1/n_1 \\ X_2/n_2 \end{pmatrix} - \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}\right) \to N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} p_1(1-p_1) & 0 \\ 0 & p_2(1-p_2) \end{pmatrix}\right) \text{ and}$$

$$n\hat{\Sigma}_n = \begin{pmatrix} \hat{p}_1(1-\hat{p}_1) & 0 \\ 0 & \hat{p}_2(1-\hat{p}_2) \end{pmatrix} = \begin{pmatrix} (X_1/n_1)(1-X_1/n_1) & 0 \\ 0 & (X_2/n_2)(1-X_2/n_2) \end{pmatrix} \xrightarrow{p} \begin{pmatrix} p_1(1-p_1) & 0 \\ 0 & p_2(1-p_2) \end{pmatrix}.$$

By Theorem 10.9, $W = \left[R\hat{P} - r\right]' \left[R\hat{\Sigma}_n R'\right]^{-1} \left[R\hat{P} - r\right] \sim \chi_1^2$, where

$R = [1 \ \ -1]$, $\hat{p} = \left[\hat{P}_1 \ \ \hat{P}_2\right]'$, $\hat{p}_i = x_i / n_i$, and $r = 0$.

b) The test is conducted by rejecting $H_0$ iff $w \in \left[\chi^2_{1;\alpha}, \infty\right)$.

$$w = (-.28366)\left((1 \ \ -1)\begin{pmatrix} .24395/45 & 0 \\ 0 & .20761/34 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)^{-1} (-.28366) = 6.98.$$

Since $w \in [2.706, \infty)$, where $\chi^2_{1;\,.10} = 2.706$, reject $H_0$.

c) The GLR statistic, given $H_0 : p_1 = p_2$ versus $H_a = p_1 \neq p_2$, is

$$\lambda(\mathbf{x}) = \frac{L(\hat{p}_0; x_1) L(\hat{p}_0; x_2)}{L(\hat{p}_1; x_1) L(\hat{p}_2; x_2)},$$

where the unrestricted MLE estimates are $\hat{p}_1 = x_1 / n_1, \hat{p}_2 = x_2 / n_2$, and the restricted MLE estimate is $\hat{p}_0 = (x_1 + x_2)/(n_1 + n_2)$.

Rewriting the GLR yields

$$\lambda(\mathbf{x}) = \frac{\binom{n_1}{x_1}\hat{p}_0^{x_1}(1-\hat{p}_0)^{n_1-x_1}\binom{n_2}{x_2}\hat{p}_0^{x_2}(1-\hat{p}_0)^{n_2-x_2}}{\binom{n_1}{x_1}\hat{p}_1^{x_1}(1-\hat{p}_1)^{n_1-x_1}\binom{n_2}{x_2}\hat{p}_2^{x_2}(1-\hat{p}_2)^{n_2-x_2}}.$$

This particular GLR statistic does not simplify easily. Theorem 10.5 can be used to construct an asymptotic size $\alpha$ GLR test based on

$-2\ln(\lambda(\mathbf{x})) \overset{a}{\sim} \chi_1^2$ (under $H_0$).

Since, $-2\ln(\lambda(\mathbf{x})) = -2\ln\left(\frac{2.2568\times 10^{-24}}{5.56992\times 10^{-23}}\right) = 6.412 > \chi^2_{1;\,.10} = 2.706,$ reject $H_0$.

19. In each of these cases one could follow the basic principles presented in Chapter 9 to attempt to define UMP or UMPU tests of the hypotheses for the finite sample case. We will instead use the GLR for the finite sample case, as well as for asymptotic cases. An LM or Wald test approach might also be pursued.

a) $L(p;x) = \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,2,\ldots,n\}}(x).$

1) $H_0: p = p_0$ versus $H_a: p \neq p_0$ (two sided)

$$\lambda(\mathbf{x}) = \frac{\binom{n}{x} p_0^x (1-p_0)^{n-x} I_{\{0,1,\ldots,n\}}(x)}{\max\limits_{p\in[0,1]} \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,\ldots,n\}}(x)} = \frac{p_0^{n\bar{x}}(1-p_0)^{n(1-\bar{x})}}{(\bar{x})^{n\bar{x}}(1-\bar{x})^{n(1-\bar{x})}}$$ (see Example 10.7, p. 606)

The GLR, $\lambda(x)$, is monotonically decreasing for increasing $\bar{x} > p_0$ or decreasing $\bar{x} < p_0$, with $\lambda(x) = 1$ being its maximum value when $\bar{x} = p_0$. Size $\alpha$ tests of $H_0$ could be constructed by finding a value of $c < 1$ such that $p(\lambda(x) \leq c) = p(\bar{x} \leq c_\ell \text{ or } \bar{x} \geq c_\mu) = \alpha$ when $p = p_0$. One can argue further that it might be possible to define a UMPU test of $H_0$; see Theorem 9.9, p. 571.

Based on Theorem 10.5, $-2\ln\lambda(X) \xrightarrow{d} \chi_1^2$. Thus "reject $H_0$ iff $-2\ln\lambda(x) \geq \chi^2_{1;\alpha}$" is an asymptotically valid size $\alpha$ test of $H_0: p = p_0$ versus $H_a: p \neq p_0$.

2) $H_0: p \leq p_0$ versus $p > p_0$ (one-sided).

The approach is analogous to Example 10.7, with .20 being replaced by $p_0$, leading to a critical region of the form

$c_r^{GLR} = \{x : \lambda(x) \leq c\} = \{x : \bar{x} \geq c_\mu\}$, where $c_\mu$, and hence $c$, is chosen so that $P\left(x \in c_r^{GLR}\right) = \alpha$ when $p = p_0$. One can argue further from Theorem 10.3 that this test will also be UMP. The case $H_0 : p \geq p_0$ versus $H_a : p < p_0$ can be handled analogously, leading to a critical region defined by $\bar{x}$ values as $\{x : \bar{x} \leq c_\ell\}$.

Asymptotically valid tests can be based on the fact that $\bar{X} \overset{a}{\sim} N\left(p_0, n^{-1} p_0 (1 - p_0)\right)$ if $p = p_0$, so that $Z = \frac{\bar{X} - p_0}{\left[n^{-1} p_0 (1 - p_0)\right]^{1/2}} \overset{a}{\sim} N(0,1)$. Then "reject $H_0$ *iff* $z \geq z_\alpha$" is an asymptotically valid size $\alpha$ test of $H_0 : p \leq p_0$ versus $H_a : p > p_0$.

b) $$L(\lambda; \mathbf{x}) = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \prod_{i=1}^n I_{\{0,1,2,\ldots\}}(x_i).$$

1) $H_0 : \lambda = \lambda_0$ versus $\lambda \neq \lambda_0$. The GLR is

$$\lambda(\mathbf{x}) = \frac{e^{-n\lambda_0} \lambda_0^{\sum_{i=1}^n x_i}}{e^{-n\bar{x}} \bar{x}^{\sum_{i=1}^n x_i}} = e^{n(\bar{x} - \lambda_0)} \left(\frac{\lambda_0}{\bar{x}}\right)^{n\bar{x}},$$

which has a minimum at $\bar{x} = \lambda_0$, and is monotonically decreasing as $\bar{x}$ increases above $\lambda_0$ or decreases below $\lambda_0$, straightforwardly shown by taking the derivative of $\ln \lambda(x)$ with respect to $\bar{x}$. Therefore, $\lambda(\mathbf{x}) \leq c$ iff $\bar{x} \leq c_\ell$ or $\bar{x} \geq c_\mu$ for appropriate choices of $c_\ell$ and $c_\mu$. Thus, an $\alpha$-size is found when $c_\ell$ and $c_\mu$ for appropriate choice of $c_\ell$ and $c_\mu$ are chosen so that $P(x \in c_r) = \alpha$, where $c_r = \{x : \bar{x} \leq c_\ell \text{ or } \bar{x} \geq c_\mu\}$.

Based on Theorem 10.5, $-2 \ln \lambda(X) \xrightarrow{d} \chi_1^2$. Thus, "reject $H_0$ iff $-2 \ln \lambda(x) \geq \chi^2_{1;\alpha}$" $s$ an asymptotically valid size $\alpha$ test of $H_0 : \lambda = \lambda_0$ versus $H_a : \lambda \neq 0$.

2) $H_0 : \lambda \leq \lambda_0$ versus $\lambda > \lambda_0$. An approach that is similar to Example 10.7 in logic, but uses the GLR defined above, leads to a critical region of the form

$C_r^{GLR} = \{\mathbf{x} : \lambda(\mathbf{x}) \leq c\} = \{\mathbf{x} : \bar{x} \geq c_\mu\}$, where $c_\mu$, and hence $c$, is chosen so that $p\left(x \in c_r^{GLR}\right) = \alpha$ when $\lambda = \lambda_0$. One could argue further from Theorem 10.3 that this test will also be UMP The case $H_0 : \lambda \geq \lambda_0$ versus $H_a : \lambda < \lambda_0$ can be handled analogously, leading to a critical region defined by $\bar{x}$ values as $\{\mathbf{x} : \bar{x} \leq c_\ell\}$.

Asymptotically valid tests can be based on the fact that $\bar{X} \overset{a}{\sim} N\left(\lambda_0, \lambda_0 / n\right)$ if $\lambda = \lambda_0$, so that $Z = \frac{\bar{X} - \lambda_0}{(\lambda_0/n)^{1/2}} \overset{a}{\sim} N\left(0,1\right)$. Then "reject $H_0$ iff $z \geq z_\alpha$" is an asymptotically valid size $\alpha$ test of $H_0 : \lambda \leq \lambda_0$ versus $H_a : \lambda > \lambda_0$.

21. It was shown in the proof of Theorem 10.14 that $F\left(Z;\theta\right) \sim \text{Uniform}\left(0,1\right)$. Using a change of variables or MGF approach, it follows that $1 - F\left(Z;\theta\right) \sim \text{Uniform}\left(0,1\right)$. Again following the proof of Theorem 10.14, a change of variables approach shows that $Y = -\ln\left(1 - F\left(Z;\theta\right)\right) \sim \text{Exponential}\left(1\right)$. Then $-2\sum_{i=1}^{n} \ln\left(1 - F\left(X_i;\theta\right)\right)$ is two times the sum of *n* independent Exponential(1) random variables, which as a $\chi^2$ distribution with 2*n* degrees of freedom. Since $Q = q\left(X;\theta\right) = -2\sum_{i=1}^{n} \ln\left(1 - F\left(X_i;\theta\right)\right) \sim \chi^2_{2n} \forall \theta \in \Omega$, it follows that *Q* is a pivotal quantity for $\theta$.

23. a) If we use Problem 10.21

$$\begin{aligned} -2\sum_{i=1}^{n} \ln\left[1 - F\left(x_i;\theta\right)\right] &= -2\sum_{i=1}^{n} \ln\left[1 - \int_0^{x_i} \theta\left(1+s\right)^{-(\theta+1)} I_{(0,\infty)}\left(s\right) ds\right] \\ &= -2\sum_{i=1}^{n} \ln\left[1 - \left(1 - \left(1 + x_i\right)^{-\theta}\right)\right] \quad x_i > 0 \\ &= 2\theta\sum_{i=1}^{n} \ln\left(1 + x_i\right) \sim \chi^2_{2n}. \end{aligned}$$

is a pivotal quantity for $\theta$.

b) A level-$\gamma = 1 - \alpha$ confidence interval for $\theta$ based on *n* iid observations from $f\left(x;\theta\right)$ is

$$\frac{\chi^2_{2n;1-\alpha/2}}{2\sum_{i=1}^{n} \ln\left(1 + x_i\right)} < \theta < \frac{\chi^2_{2n;\alpha/2}}{2\sum_{i=1}^{n} \ln\left(1 + x_i\right)}.$$

c) Let $\gamma = .90, n = 100$.

Using the approximation for critical points of the $\chi^2$ distribution indicated in the note, we find

$$\chi^2_{200;\,.05} = 200\left[1-\left(\frac{2}{9(200)}\right)+1.645\left(\frac{2}{9(200)}\right)^{1/2}\right]^3 = 233.996,$$

$$\chi^2_{200;\,.95} = 200\left[-1\left(\frac{2}{9(200)}\right)-1.645\left(\frac{2}{9(200)}\right)^{1/2}\right]^3 = 168.276.$$

(Note the actual values of these critical values are 233.994 and 168.279.)

Then using the result in b), the confidence interval outcome can be calculated as

$$\frac{168.276}{2(40.54)} < \theta < \frac{233.996}{2(40.54)},$$
$$\Leftrightarrow 2.075 < \theta < 2.886.$$

# Appendix A – Answer Key

1. a) $S_1 = \{x: x$ is a senior citizen receiving social security payments in the United States$\}$.

   b) $S_2 = \{x: x = 10^n, n=1,2,3,...\}$.

   c) $S_3 = \{\gamma: \gamma=y-x, x, y \in \{1, 2, 3, 4, 5, 6\}\}$.

   d) $S_4 = \{(x_1, x_2): x_1 \in R, x_2 = e^{x_i}\}$.

3. a) $S$ is finite

   b) $S$ is uncountably infinite

   c) $S$ is finite

   d) $S$ is countably infinite

5. a) $\cup_{i\in I} A_i = A_i \cup A_3 \cup A_4 = [-5,1) \cup (2,5]$.

   b) $\cup_{i=1}^{4} = A_1 \cup A_2 \cup A_3 \cup A_4 = [-5,1) \cup (1,5]$.

   c) $A_1 \cap A_2 = \varnothing$.

   d) $A_4 - A_1 = [-5,-2)$.

   e) $\overline{A}_4 = (-2,5]$.

7. Both satisfy Definition A.17, and so are functions.

   a) $D(f) = A$, $R(f) = B$

   b) $D(f) = A$, $R(f) = B$

9. a) Yes

   b) Yes

   c) Yes

   d) $f(2) = 8$; $f^{-1}(5) = 1$.

   e) $D(f)=A$; $R(f) = (2,\infty)$.

11. $f$: $A \to R$

$$f = \{(A,y) : y = (x_2 - x_1)(y_2 - y_1), A = [x_1,x_2] \times [y_1,y_2], x_2 \geq x_1, y_2 \geq y_1\},$$
$$R(f) = \{y : y = (x_2 - x_1)(y_2 - y_1), x_2 \geq x_1, y_2 \geq y_1\},$$
$$D(f) = \{A : A = [x_1,x_2] \times [y_1,y_2], x_2 \geq x_1, y_2 \geq y_1\}.$$

13. a) $\binom{15}{3} = 455.$

b) $\binom{14}{2} = 91.$

15. a) $\sum_{x \in A_2} x \;\; = 1+2+3+4+5 = 15.$

b) $\sum_{i \in A_2} y_i - \sum_{i=1}^{5} y_i = \sum_{i=1}^{5} i^2 = 55.$

c) $\sum_{x_1 \in A_1} \sum_{x_2 \in A_2} \left(\frac{1}{2}\right)^{x_1} x_2^2 = 55 \sum_{i=1}^{\infty} (1/2)^i = 55.$

d) $\sum_{x \in A_1 - A_2} (1/3)^x = \sum_{i=6}^{\infty} (1/3)^i = 0.0021.$

e) $\sum \sum_{(x_1,x_2) \in B} (x_1 + x_2) = 90.$