

1. Aufgabe

- a. Welchen Einfluss hat die Größe der Stichprobe auf die Breite des Konfidenzintervalls? Wie verändert sich diese Breite beim Wachsen bzw. Fallen der Größe der Stichprobe?

$$\text{Es ist } G^2 \geq \frac{\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2}{n}, \text{ also } G \geq \frac{\left|\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right|}{\sqrt{n}}$$

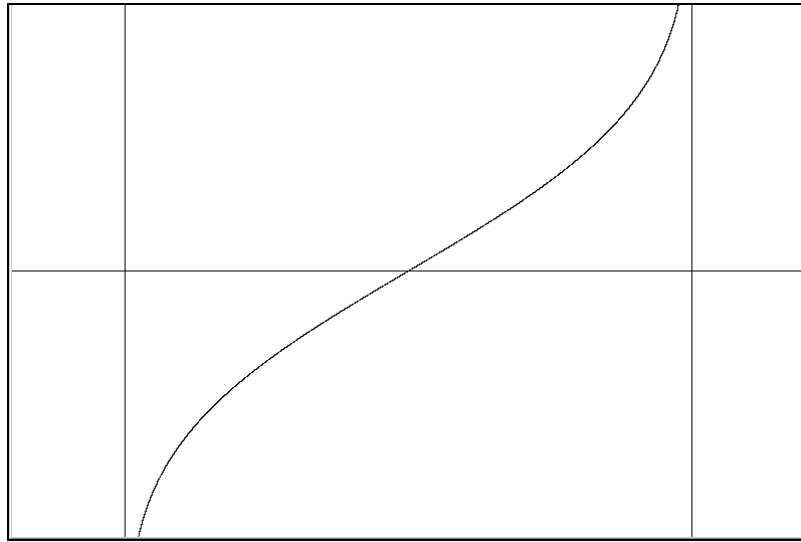
d.h. Das Konfidenzintervall kann immer schmaler werden, wenn die Größe n der Stichprobe wächst. Denn dann wird auch die Wurzel größer und der Kehrwert wird kleiner.

$$\text{Es ist weiter } n \geq \frac{\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2}{G^2}$$

Das bedeutet, dass n immer kleiner werden kann, wenn die Breite G des Konfidenzintervalls wächst und dass andererseits eine geringer werdende Breite eine wachsende Stichprobengröße verlangt. Das Wachstum ist quadratisch bei umgekehrter Proportionalität.

- b. Welchen Einfluss hat die Größe des Konfidenzniveaus auf die Breite des Konfidenzintervalls? Wie verändert sich diese Breite beim Wachsen bzw. Fallen der Größe des Konfidenzniveaus?

Graph von Φ^{-1}



Es war $G \geq \frac{\left| \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \right|}{\sqrt{n}}$. Der Graph von Φ^{-1} zeigt, dass bei kleiner werdendem Konfidenzniveau ($\alpha \rightarrow 0$) die notwendige Breite des Konfidenzintervalls sehr schnell wächst. Falls α wächst, falls also die geforderte Konfidenzwahrscheinlichkeit kleiner wird, kann auch das Konfidenzintervall kleiner werden. Bei $\alpha = 1$ schließlich (das entspricht der geforderten Konfidenzwahrscheinlichkeit 0) kann dieses Intervall sogar beliebig klein werden.

2. Aufgabe

Wie schon in den vergangenen Kapiteln interessieren wir uns über den Anteil fehlerhaft produzierter Warenstücke in einer Massenproduktion.

- a. Bei der Überprüfung einer Stichprobe von 1000 Elementen findet man 93 defekte Teile. Es soll nun zum Konfidenzniveau 95 % ein Konfidenzintervall für die unbekannte Wahrscheinlichkeit p ermittelt werden, mit der ein fehlerhaftes Teil produziert wird. Arbeiten Sie mit der Formel von Satz 25.9 und vergewissern Sie sich, dass die Voraussetzungen dieses Satzes erfüllt sind.

Natürlich ist die Zufallsvariable H_{1000} = Anzahl der defekten Teile binomial verteilt. Es ist $H_{1000} = 93 > 30$ und $1000 - 93 > 30$. Sei nun $\alpha = 0,05$. Dann ist $c = \Phi^{-1}(0,975) = 1,96$ und wir erhalten als Konfidenzintervall für p zur Konfidenzwahrscheinlichkeit 95% das Intervall:

$$[0,093 - 0,018, 0,093 + 0,018]$$

- b. Bei der Überprüfung einer Stichprobe von 2000 Elementen findet man 186 defekte Teile. Es soll wieder zum Konfidenzniveau 95 % ein Konfidenzintervall für die unbekannte Wahrscheinlichkeit p ermittelt werden, mit der ein fehlerhaftes Teil produziert wird. Arbeiten Sie auch hier mit der Formel von Satz 25.9.

Auch hier sind die Voraussetzungen von Satz 25.9 erfüllt. Wieder ist $c = 1,96$ und wir erhalten als Konfidenzintervall für p zur Konfidenzwahrscheinlichkeit 95% das Intervall:

$$[0,093 - 0,013, 0,093 + 0,013]$$

- c. Die Intervalle in Aufgabe 2 a und 2 b haben beide denselben Mittelpunkt: $93/1000 = 186/2000 = 0,093 = 9,3 \%$. Erklären Sie die unterschiedliche Breite der beiden Intervalle.

Die Größe der Stichprobe ist in der Aufgabe 2 b gewachsen, alle anderen Parameter sind gleich geblieben, von daher ergibt sich in Aufgabe 2 b ein kleineres Konfidenzintervall, d.h. die Schätzung für die unbekannte Wahrscheinlichkeit p ist genauer.

- d. Ermitteln Sie die erforderliche Stichprobengröße, die zum Konfidenzniveau 95% mit den in (a) und (b) verwendeten Verfahren Konfidenzintervalle liefert, deren Längen nicht größer als 0,02 sind.

Wir nehmen zunächst immer an: $H_n = 0,093 \cdot n$. In diesem Falle muss gelten:

$$\frac{c}{n} \sqrt{\frac{H_n(n - H_n)}{n}} = \frac{1,96}{n} \sqrt{0,093 \cdot 0,907 \cdot n} = \frac{0,57}{\sqrt{n}} \leq 0,01$$

$$\text{Also } \sqrt{n} \geq \frac{0,57}{0,01} = 57, \text{ also } n \geq 3249$$

Falls wir mit Satz 25.10 arbeiten und die Annahme für die empirische Häufigkeit weglassen, erhalten wir die Anforderung:

$$n \geq \frac{c^2}{G^2} = \frac{3,8416}{0,0004} = 9604$$

3. Aufgabe

2005 wurden in Deutschland 334399 Mädchen und 351721 Jungen geboren. Bestimmen Sie auf Grund dieser Daten zum Konfidenzniveau 99 % ein Konfidenzintervall für die Wahrscheinlichkeit einer Mädchengeburt. Arbeiten Sie wieder mit der Formel von Satz 25.9.

Hier ist $\alpha = 0,01$ und $c = \Phi^{-1}(0,995) = 2,575$, $n = 334399 + 351721 = 686120$ und $H_n = 334399$. Es ist weiter:

$$\frac{H_n}{n} = \frac{334399}{686120} = 0,487 \quad \text{und} \quad \frac{c}{n} \sqrt{\frac{H_n(n - H_n)}{n}} = 0,00155$$

Das bedeutet, unser Konfidenzintervall lautet:

$$[0,487 - 0,00155, 0,487 + 0,00155]$$

Es hat die (sehr geringe) Breite von 0,0031

4. Aufgabe

Es sind Parlamentswahlen und ein Meinungsforschungsinstitut möchte den prozentualen Stimmenanteil der Partei PDS (**P**artei **d**er **S**tatistiker) prognostizieren.

- a. Wie viele zufällig ausgewählte Wahlberechtigte müssen mindestens befragt werden, um für den prozentualen Stimmenanteil ein Konfidenzintervall zum Niveau 0,96 zu erhalten, dessen Länge höchstens 2 Prozent beträgt.

$$\begin{aligned} \text{Es war } n &\geq \frac{\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2}{G^2} \geq \frac{(\Phi^{-1}(0,98))^2}{0,0004} = \frac{2,055^2}{0,0004} \\ &\geq 10558 \end{aligned}$$

- b. Wie viele zufällig ausgewählte Wahlberechtigte müssen mindestens befragt werden, um für den prozentualen Stimmenanteil ein Konfidenzintervall zum Niveau 0,98 zu erhalten, dessen Länge höchstens 2 Prozent beträgt.

$$n \geq \frac{(\Phi^{-1}(0,99))^2}{0,0004} = \frac{2,325^2}{0,0004} \geq 13515$$

- c. Das Meinungsforschungsinstitut befragt 12000 Personen. Davon behaupten 2160, dass sie die Partei der Statistiker wählen werden. Bestimmen Sie sowohl für das Konfidenzniveau 96 % als auch für das Konfidenzniveau 98 % die zugehörigen Konfidenzintervalle für die Wahrscheinlichkeit, dass diese Partei gewählt wird. Arbeiten Sie mit der Formel von Satz 25.9.

Zunächst ist $\alpha = 0,04$ und $c = \Phi^{-1}(0,98) = 2,055$, $n = 12000$ und $H_n = 2160$. Es ist weiter:

$$\frac{H_n}{n} = \frac{2160}{12000} = 0,18 \quad \text{und} \quad \frac{c}{n} \sqrt{\frac{H_n(n - H_n)}{n}} = 0,0072$$

Das bedeutet, unser Konfidenzintervall lautet:

$$[0,18 - 0,0072, 0,18 + 0,0072]$$

Es hat die Breite von 0,0144

Im zweiten Teil der Aufgabenstellung ist $\alpha = 0,02$ und $c = \Phi^{-1}(0,99) = 2,325$ und wieder $n = 12000$ und $H_n = 2160$. Jetzt muss G größer werden. Es ist:

$$\frac{c}{n} \sqrt{\frac{H_n(n - H_n)}{n}} = 0,0082$$

Das bedeutet, unser Konfidenzintervall lautet nun:

$$[0,18 - 0,0082, 0,18 + 0,0082]$$

Es hat die Breite von 0,0164

5. Aufgabe

- a. Nehmen Sie an, Sie wollten für eine Messreihe, in der immer wieder dasselbe Bernoulliexperiment durchgeführt wird, ein Konfidenzintervall für die zu suchende unbekannte Wahrscheinlichkeit p finden, mit der ein bestimmtes Ergebnis eintritt. Sie möchten absolute Sicherheit haben, d.h. das Konfidenzniveau soll 100 % betragen. Wie muss das Konfidenzintervall aussehen?

Das Konfidenzintervall muss von 0 bis 1 gehen, eigentlich (wenn die empirischen Häufigkeiten nicht = 0 oder = n sind) müsste es von der Art $]0, 1[$ sein, sonst erhalten wir $[0, 0]$ oder $[1, 1]$

- b. Warum muss Ihre Antwort im Widerspruch zu der im Anhang 2 abgedruckten Tafel von Funktionswerten der Normalverteilung Φ stehen und wie erklären Sie diesen Widerspruch?

Falls $\alpha = 0$ ist, ist $c = \Phi^{-1}(1)$ nicht definiert bzw. $= \infty$ und wir erhalten für empirische Häufigkeiten, die nicht 0 oder nicht = n sind, ein unendlich großes Konfidenzintervall.

Für diese Grenzsituation $\alpha = 0$ ist die Formel aus Satz 25.9 nicht geeignet, da dann c nicht definiert ist.