AUFGABENTYPEN

1. Bekannt ist die Entscheidungsregel, d.h. K und $\overline{K}$;
   zu berechnen ist das Risiko 1.Art (bzw. 2. Art).

2. Bekannt ist die Irrtumswahrscheinlichkeit α;
   zu berechnen ist der Annahme- und Ablehnungsbereich, also die Entscheidungsregel.

3. Bekannt ist die Irrtumswahrscheinlichkeit α und ein Versuchsergebnis.
   Zu untersuchen ist, ob die Nullhypothese aufgrund des Versuchsergebnisses angenommen oder abgelehnt wird (dazu bestimmt man K und $\overline{K}$ zur vorgegebenen Irrtumswahrscheinlichkeit α.

Beispiele zur Bestimmung von K und $\overline{K}$

Beispiel 1: (einseitiger Test über p)

Ein Lieferant von Saatkartoffeln teilt seinem Abnehmer mit, der Virusbefall der gelieferten Kartoffeln (ein Eisenbahnwagen) sei genau 20%. Er vereinbart mit dem Käufer einen Preisnachlass, falls sich bei einem Test der Virusbefall größer als 20% erweisen sollte. Wegen der hohen Laborkosten kommen beide überein, der Sendung nur eine kleine Stichprobe im Umfang von 20 zufällig zu entnehmen und untersuchen zu lassen. Der Lieferant nimmt ein Risiko von maximal 10% auf sich, zu Unrecht einen Preisnachlass gewähren zu müssen. Auf welchen Annahmebereich werden sich die beiden einigen?

Bemerkung:
Es handelt sich hier zwar um ein Ziehen ohne Zurücklegen, d.h. also um eine hypergeometrische Verteilung. Da aber N (die Menge der Kartoffeln) sehr groß ist, können wir diese durch die Binomialverteilung ersetzen.

p=WS, eine virusbefallene Kartoffel zu entnehmen

$$\boxed{H_0\text{: } p=0{,}2 \quad \text{oder} \quad H_1\text{: } p>0{,}2}$$

X=Zufallsvariable für die Anzahl der virusbefallenen Kartoffeln mit den Werten M={0,1,2,3,4,...,20}

Vorüberlegung:
Für kleine Zahlen aus M wird man H_0 annehmen, für große Zahlen ablehnen.

Ansatz für K: K={0,1,2,3,...,c-1} $\Rightarrow$ $\overline{K}$={c, c+1, c+2,...,20}

Entscheidungsregel:

$X \in K \ \rightarrow$ Annahme von H_0

$X \in \overline{K} \ \rightarrow$ Ablehnung von H_0

Fehler 1. Art: Preisnachlass zu Unrecht, falls H_0 abgelehnt wird, obwohl H_0 wahr ist.

Risiko 1. Art: $P(X \in \{c, c+1, c+2, \ldots, 20\}) = \sum_{i=c}^{20} \binom{20}{i} \cdot 0{,}2^i \cdot 0{,}8^{20-i}$

$$\sum_{i=c}^{20} \binom{20}{i} \cdot 0{,}2^i \cdot 0{,}8^{20-i} \overset{!}{\leq} 0{,}1 \qquad \text{c möglichst klein!}$$

$$1 - \sum_{k=0}^{c-1} \binom{20}{k} \cdot 0{,}2^k \cdot 0{,}8^{20-k} \leq 0{,}1$$

$$P(X \in K) = 1 - \sum_{k=0}^{c-1} \binom{20}{k} \cdot 0{,}2^k \cdot 0{,}8^{20-k} > 0{,}9$$

$$\underset{CAS}{\overset{GTR}{\Longrightarrow}} \quad c-1 = 6 \quad \Rightarrow \quad c = 7$$

Also $K_{0,2} = \{0,1,2,3,4,5,6\}$, $\overline{K}_{0,2} = \{7,8,\ldots,20\}$

Problem: Der Lieferant behauptet, der Virusbefall sei <u>höchstens</u> 20%.

$$\boxed{H_0\!: p \leq 0{,}2 \quad \text{oder} \quad H_1\!: p > 0{,}2}$$

p=0,2: $K_{0,2} = \{0,1,2,3,4,5,6\}$, $\overline{K}_{0,2} = \{7,8,\ldots,20\}$

p<0,2: z.B. p=0,1:

Bestimmung von $K_{0,1}$ und $\overline{K}_{0,1}$:

$K_{0,1} = \{0,1,2,3,\ldots,c-1\}$ und $\overline{K}_{0,1} = \{c, c+1, c+2, \ldots, 20\}$

$$P(X \in \{c, c+1, c+2, \ldots, 20\}) = \sum_{i=c}^{20} \binom{20}{i} \cdot 0{,}1^i \cdot 0{,}9^{20-i}$$

$$\sum_{i=c}^{20} \binom{20}{i} \cdot 0{,}1^i \cdot 0{,}9^{20-i} \overset{!}{\leq} 0{,}1 \qquad \text{c möglichst klein!}$$

$$1 - \sum_{k=0}^{c-1} \binom{20}{k} \cdot 0{,}1^k \cdot 0{,}9^{20-k} \leq 0{,}1$$

$$P(X \in K) = 1 - \sum_{k=0}^{c-1} \binom{20}{k} \cdot 0{,}1^k \cdot 0{,}9^{20-k} > 0{,}9$$

$$\underset{CAS}{\overset{GTR}{\Longrightarrow}} \quad c-1 = 4 \quad \Rightarrow \quad c = 5$$

Also $K_{0,2} = \{0,1,2,3,4\}$, $\overline{K}_{0,2} = \{5,6,7,8,\ldots,20\}$

Allgemein:
Für alle $p<0,2$ erhält man zur Irrtumswahrscheinlichkeit $\alpha=0,1$ einen klieneren Annahmebereich als für $p=0,2$., denn falls $p<0,2$ gilt, wird X noch eher einen kleineren Wert annehmen als bei $p=0,2$, d.h. X wird mit noch größerer WS in dem Bereich der einfachen Nulllhypothese H_0: $p=0,2$ fallen.

Falls man also zur Irrtumswahrscheinlichkeit $\alpha=0,1$ einen gemeinsamen Annahmebereich sucht, der für alle $p\le 0,2$ verwendbar ist, so kann man $K_{0,2}$ als gemeinsamen Annahmebereich wählen.

Sei K_p, $p\le 0,2$, der Annahmebereich (zu p) und zur Irrtumswahrscheinlichkeit $\alpha=0,1$, d.h. es gilt $P(X\in\overline{K_p})\le\alpha=0,1$

$p\le 0,2 \Rightarrow \overline{K_p}\supset\overline{K_{0,2}} \Rightarrow P(X\in\overline{K_{0,2}})\le P(X\in\overline{K_{p<0,2}})\le\alpha=0,1$

d.h. also:
$K_{0,2}$ ist ein Annahmebereich, der für alle $p\le 0,2$ eine Irrtumswahrscheinlichkeit von $\alpha=0,1$ garantiert.

Allgemein:

$H_0: \ p\le p_0 \quad (p\ge p_0)$ oder $H_1: \ p>p_0 \quad (p<p_0)$

Zusammenges. Hyp Zusammenges. Hyp

Da X für $p<p_0 \quad (p\ge p_0)$ mit noch größerer WS in den Annahmebereich der einfachen Hypothese $H_0{:}p=p_0$ fällt, wählt man als Annahmebereich für $H_0{:} \quad p\le p_0 \quad (p\ge p_0)$ den Annahmebereich der einfachen Nullhypothese $H_0{:}p=p_0$.

Beispiel 2 (zweiseitiger Test über p):

Ein Lieferant von Saatkartoffeln teilt seinem Abnehmer mit, der Virusbefall der Kartoffeln sei 20%. Er vereinbart mit dem Käufer einen Preisnachlass, falls sich der Virusbefall bei einem Test deutlich größer als 20% herausstellen sollte, und einen Preisaufschlag, wenn das Testergebnis für einen kleineren Anteil als 20% spricht. Beide kommen überein, 25 Kartoffeln untersuchen zu lassen. Der Lieferant ist bereit, ein Risiko bis maximal 20% einzugehen, vom vereinbarten Normalpreis abweichen zu müssen. Welcher Bereich eignet sich für die Annahme der Nullhypothese $p=0,2$?

Bemerkung:
N ist groß, deshalb kann man mit der Binomialverteilung rechnen.
p=WS eine virusbefallene Kartoffel zu nehmen

$\boxed{H_0{:}\ p=0,2 \quad \text{oder } H_1{:}p\ne 0,2}$

X ist die Zufallsvariable für die Anzahl der virusbefallenen Kartoffeln mit Werten aus M={0,1,2,...,25}

Vorüberlegung:

Man wird die Nullhypothese H_0: p=0,2 für zu kleine oder für zu große Werte von X ablehnen.

Ansatz für K und $\overline{K}$:

$\overline{K}_1 = \{0,1,2,...,c\}$

$\overline{K}_2 = \{d,...,25\}$

$\overline{K} = \overline{K}_1 \cup \overline{K}_2 = \{0,1,2,...,c\} \cup \{d,...,25\}$ Ablehnungsbereich

K={c+1,..., d-1} Annahmebereich

Preisänderung zu Unrecht ist ein Fehler erster Art:

H_0:p=0,2 ist wahr

Risiko 1. Art = $P(X \in \overline{K}) = \underbrace{\sum_{i=0}^{c} \binom{25}{i} \cdot 0{,}2^i \cdot 0{,}8^{25-i}}_{\leq 10\%} + \underbrace{\sum_{i=d}^{25} \binom{25}{i} \cdot 0{,}2^i \cdot 0{,}8^{25-i}}_{\leq 10\%} \overset{!}{\leq} 0{,}2 = 20\%$

Das Risiko erster Art von maximal 20% wird gleichmäßig aufgeteilt in

Risiko für Preisnachlass $\leq$ 10%

Risiko für Preisaufschlag $\leq$ 10%

Risiko für Preisaufschlag = $\sum_{i=0}^{c} \binom{25}{i} \cdot 0{,}2^i \cdot 0{,}8^{25-i} \overset{!}{\leq} 10\% = 0{,}1$ „c möglichst groß“

$\Rightarrow c=2 \quad \Rightarrow \overline{K}_1 = \{0,1,2\}$

Risiko für Preisnachlass = $\sum_{i=d}^{25} \binom{25}{i} \cdot 0{,}2^i \cdot 0{,}8^{25-i} \overset{!}{\leq} 10\% = 0{,}1$ „d möglichst klein“

$$1 - \sum_{i=0}^{d-1} \binom{25}{i} \cdot 0{,}2^i \cdot 0{,}8^{25-i} \overset{!}{\leq} 10\% = 0{,}1$$

$$\sum_{i=0}^{d-1} \binom{25}{i} \cdot 0{,}2^i \cdot 0{,}8^{25-i} \geq 0{,}9$$

$\Rightarrow d-1 = 8 \quad \Rightarrow \overline{K}_2 = \{9,10,...,25\}$

also $\mathbf{\overline{K} = \{0,1,2\} \cup \{9,10,...,25\}}$ und **K = {3,4,...8}**

Bemerkungen zu K und $\overline{K}$:

1. Man hätte das Risiko 1. Art von maximal 20% auch unregelmäßig aufteilen können (z.B. 5% und 15%) und dann mithilfe dieser Aufteilung K und $\overline{K}$ bestimmen können.

2. Man könnte auch mithilfe des Erwartungswertes von X (unter der Voraussetzung, dass H_0 wahr ist) den Ablehnungsbereich und den Annahmebereich bestimmen:

Da N groß ist, kann man den Erwartungswert von X wie bei der Binomialverteilung berechnen:

$E(X) = 25 \cdot 0{,}2 = 5$

$K = \{5-c, \ldots, 4, 5, 6, \ldots, 5+c\}$ = symmetrische Umgebung von $E(X)=5$

$= \{X \mid |X-5| \leq c\}$

$P(X \in \overline{K}) \leq 0{,}2$

$1-P(X \in K) \leq 0{,}2$

$P(X \in K) \geq 0{,}8$

$P(|x-5| \leq c) \geq 0{,}8$ c möglichst klein!

Versuch mit c = 2 :

$P(3 \leq X \leq 7) = P(X \leq 7) - P(X \leq 2) = 0{,}79265$

Versuch mnit c = 3:

$P(2 \leq X \leq 8) = P(X \leq 8) - P(X \leq 1) = 0{,}92584$

Damit gilt also: $K=\{2,3,\ldots,8\}$ und $\overline{K} = \{0,1\} \cup \{9,10,\ldots,25\}$

Allgemein – zweiseitiger Test:

H_0: $p=p_0$ oder H_1: $p \neq p_0$

X sei die Zufallsvariable für die Anzahl eines bestimmten Ereignisses mit Werten aus $M=\{0,1,2,\ldots,n\}$.

α sei die vorgegeben Irrtumswahrscheinlichkeit.

$\overline{K}_1 = \{0,1,2,\ldots,c\}$

$\overline{K}_2 = \{d,\ldots,25\}$

$\overline{K} = \overline{K}_1 \cup \overline{K}_2 = \{0,1,2,\ldots,c\} \cup \{d,\ldots,25\}$ $K=\{c+1,\ldots, d-1\}$

Zu bestimmen sind c und d so, dass gilt:

$P(X \in \overline{K}_1) = \sum_{i=0}^{c} \binom{n}{i} \cdot p_0^{\,i} \cdot (1-p_0)^{n-i} \leq \frac{\alpha}{2}$ c möglichst groß

$P(X \in \overline{K}_2) = \sum_{i=d}^{n} \binom{n}{i} \cdot p_0^{\,i} \cdot (1-p_0)^{n-i} \leq \frac{\alpha}{2}$ d möglichst klein