Aufgabe C1 Landesabitur Hessen 2012 LK

Aufgabe 1.1 (5 BE)

Die Aufgabe dreht sich um einen fiktiven Zeitungsartikel, der das gehäufte Auftreten einer Phantasiekrankheit TIN in einer deutschen Kleinstadt beschreibt.
Statistisch gesehen dürfte es in Usse, so heißt die ebenfalls frei erfundene Gemeinde, unter den Männern 5,2 TIN-Fälle geben. In Wirklichkeit tritt die Krankheit aber 12 Mal auf. Wichtig dabei zu wissen : In Usse leben 9625 Menschen, 49 % davon sind Männer.
Im Aufgabenteil 1.1 ist jetzt gefragt, wie groß die Wahrscheinlichkeit für einen Mann ist, innerhalb von 8 Jahren in Deutschland an TIN zu erkranken. Das Ergebnis p = 0,11 % wird vorgegeben, um Folgefehler in den späteren Aufgabenteilen möglichst zu vermeiden.

Ist A das Ereignis, dass ein Mann an TIN erkrankt, so ist p(A) gesucht.

$$p(A) = \frac{\text{Anzahl der möglichen Erkrankungen}}{\text{Anzahl der Männer in Usse}} = \frac{5,2}{0,49 \cdot 9625} \approx 0,001103 \approx 0,11\%$$

Dabei geht man natürlich davon aus, dass der Anteil der Erkrankungen in Usse dem Anteil der Erkrankungen in der gesamten Bundesrepublik entspricht.

Aufgabe 1.2 (5 BE)

Von den rund 82 Millionen Einwohner in Deutschland wurden in den letzten 8 Jahren 2500 zufällig ausgewählte Männer beobachtet. Die Rechnung

$$\sum_{k=0}^{3} \binom{2500}{k} \cdot 0,0011^k \cdot 0,9989^{2500-k} \approx 70,3\%$$

soll in diesem Zusammenhang erläutert werden.
Sei X die Zahl der an TIN erkrankten Männer unter den 2500 Ausgesuchten, so wird in der gegebenen Zeile $p(X \le 3)$ berechnet, die Wahrscheinlichkeit also, dass maximal 3 der 2500 Kandidaten an TIN erkranken. Dabei werden 4 Bernoulliketten aufsummiert. Zieht man Bernoulliketten zur Berechnung von Wahrscheinlichkeiten heran, sollte immer überprüft werden, ob die drei notwendigen Voraussetzungen dafür erfüllt sind :

Voraussetzung	**Sachzusammenhang**
Es gibt nur 2 mögliche Ausgänge	An TIN erkrankt oder nicht
Die Wahrscheinlichkeiten für die beiden Ausgänge ändern sich nicht	Die Wahrscheinlichkeiten bleiben auf Grund der großen Stichprobenmenge nahezu gleich p(krank) = 0,0011 p(nicht krank) = 0,9989
Die Bernoulli –Kette ist beliebig lang	Der Stichprobenumfang ist beliebig erweiterbar

Damit ist die Zufallsgröße X binomialverteilt und die obige Summenformel anwendbar.

Aufgabe 2.1 (2 BE)

Die erhöhte Zahl von TIN-Erkrankungen in Usse, 12 gegenüber der statistisch erwarteten Zahl von 5,2 Erkrankungen, nährt den Verdacht, dass in Usse ein signifikant höheres Krankheitsrisiko vorhanden ist. Diese Vermutung soll mit einem Hypothesentest gestützt werden. Dazu müssen 2 Hypothesen, die Nullhypothese und die Alternativhypothese, formuliert werden. Dies ist Inhalt diese Aufgabenteils.

Ho (Nullhypothese) :	Das Risiko, in Usse an TIN zu erkranken, beträgt 0,11 %. Dies ist die gesicherte Erkenntnis, ein bundesweit zu beobachtender, langjähriger Erfahrungswert. Die Nullhypothese beschreibt immer denn Status quo, nichts hat sich geändert (oft am Gleichheitszeichen zu erkennen. Hier **p =** 0,11 %)
H1 (Alternativhypothese) :	Das Risiko, in Usse an TIN zu erkranken ist größer als 0,11 %. Dies ist eine Vermutung, eine Behauptung, die sich auf Grund der Zahl der Krankheitsfälle aufdrängt und die jetzt getestet werden soll.

Der Grundgedanke eines Hypothesentests besteht nun darin, den α-Fehler möglichst gering zu halten. Unter dem α-Fehler versteht man den Fehler, die Nullhypothese abzulehnen, obwohl sie in Wirklichkeit zutrifft. Anders formuliert, der Alternativhypothese Glauben zu schenken, obwohl sie falsch ist. Dieser Fehler sollte immer möglichst klein gehalten werden, um nicht vorschnell einer Vermutung, einer Behauptung, einem neuen Medikament, … den Vorrang gegenüber der vertrauten alten Umgebung zu geben.

Ein großer α-Fehler in diesem Sachzusammenhang würde bedeuten, dass man nur mit einer großen Unsicherheit behaupten kann, die Zahl der Erkrankten in Usse sei signifikant höher als im übrigen Bundesgebiet. Man nennt den Betrag des α-Fehlers das Signifikanzniveau (des Tests), p(α) die Irrtumswahrscheinlichkeit.

Wie muss ein Hypothesentest nun gestaltet werden, wenn das Signifikanzniveau, die Irrtumswahrscheinlichkeit bzw. der α-Fehler nicht mehr als 5 % betragen soll ?

Aufgabe 2.2 (8 BE)

X :	Anzahl der in Usse an TIN erkrankten Männer
n =	4716, die Größe der Stichprobe (hier Zahl aller Männer in Usse)
k :	kritische Zahl. Ist die Zahl der erkrankten Männer größer als k, akzeptiert man die Alternativhypothese.
$p_{Ho}(X>k)$	ist die Wahrscheinlichkeit, die Alternativhypothese anzunehmen, obwohl Ho zutrifft. Entspricht dem α-Fehler.

Den Test nennt man rechtsseitig, weil der Ablehnungsbereich der Nullhypothese am rechten Rand liegt. Mit diesen Bezeichnungen gilt :

$$\alpha = p_{Ho}(X>k) = 1 - p_{Ho}(X \leq k) \leq 0{,}05 \qquad \text{daraus folgt}$$

$$p_{Ho}\left(X \leq k\right) \geq 0{,}95$$

Die Zufallsvariable X ist dabei B(4716 ; 0,0011) - verteilt. An Hand der mitgelieferten Tabelle lässt sich leicht ablesen

$$F_{4716;0,0011}(9) = 0{,}9609 \quad \text{und} \quad F_{4716;0,0011}(8) = 0{,}9191$$

Bei der kritischen Zahl k = 9 endet somit der Annahmebereich der Nullhypothese. Der Ablehnungsbereich von Ho beginnt bei 10 und endet bei 4716. Da in Usse 12 Erkrankungen vorliegen, kann man also von einer signifikanten Abweichung vom bundesweiten Durchschnitt sprechen.

Aufgabe 2.3 (6 BE)

In diesem Teil der Aufgabe soll davon ausgegangen werden, dass die Wahrscheinlichkeit für einen Mann in Usse, an TIN zu erkranken, tatsächlich höher ist als in der übrigen Republik, nämlich p = 0,25 %. Der Fehler 2. Art ist zu berechnen.

Zur Erinnerung : Einen Fehler 2. Art begeht man, wenn man die Nullhypothese Ho annimmt, obwohl sie falsch ist oder, was das selbe ist, wenn man die Alternativhypothese ablehnt, obwohl sie stimmt. Den Fehler 2. Art nennt man auch β - Fehler. Er ist nur berechenbar, wenn p(H1) bekannt ist, in unserem Fall

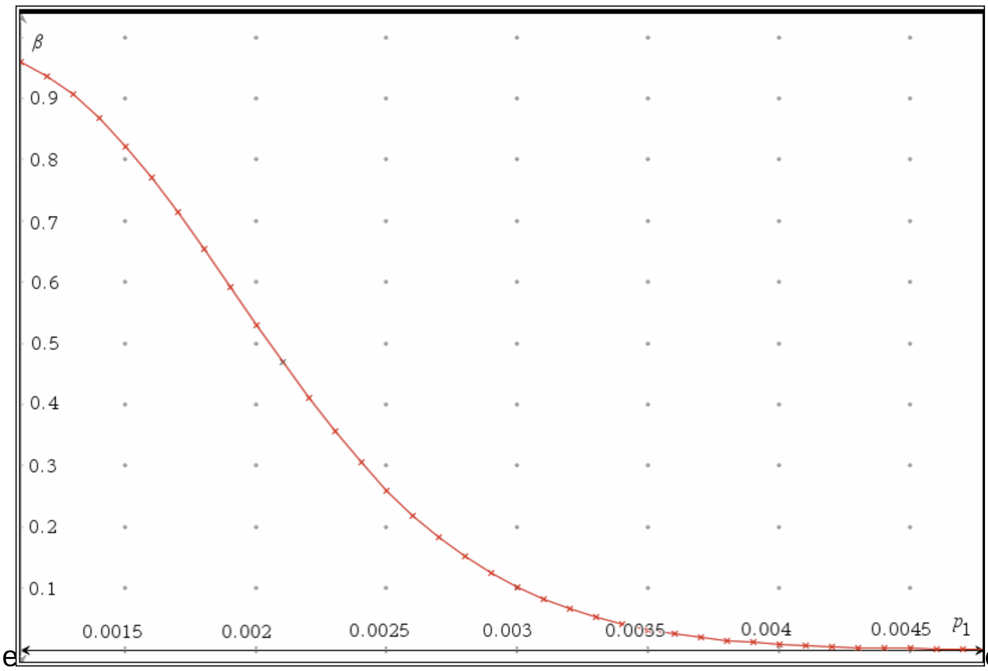
p(H1) = 0,0025 Damit wird

$$\beta = p_{H1}\left(X \leq k\right) = p_{0,0025}\left(X \leq 9\right) \approx 0{,}2609 \quad \text{laut Tabelle}$$

Ein β - Fehler von 26,09% bedeutet in unserem Zusammenhang, dass man mit einer Wahrscheinlichkeit von 26,1% davon ausgeht, dass das Risiko an TIN zu erkranken nicht größer als im restlichen Deutschland ist (0,11%), obwohl das Risiko, für Männer, im Gemeindeverband Usse an TIN zu erkranken, tatsächlich bei 0,25% liegt und damit deutlich erhöht ist.

Aufgabe 2.4 (4 BE)

Material 2 zeigt die Operationscharakteristik des Hypothesentests mit einem Signifikanzniveau von 5 %.



Der rote Graph stellt eine Verallgemeinerung der Aufgabe 2.3 dar. Für $p(H1) = 0{,}0025$ auf der x-Achse kann man den β - Fehler $\beta = 0{,}2609$ auf der y-Achse ablesen. Umgekehrt kann man bei einem vorgegebenen β - Fehler auf der x-Achse ablesen, ab welcher Wahrscheinlichkeit p(H1) dieser Fehler eintritt.

In unserem Fall soll der β - Fehler höchstens 10 % betragen. Ab einer Wahrscheinlichkeit von 0,003 ist diese Forderung erfüllt, da es sich bei der Operationscharakteristik um eine monoton fallende Kurve handelt.

