Aufgabe 1

Statistix begleitet als liebender Beziehungspartner häufig seine Freundin Statistica bei deren Lieblingsbeschäftigung, Shoppen in der City. Die langen Wartezeiten an den Kassen der Geschäfte machen den ungeduldigen Statistix jedoch regelmäßig wütend, was zum Streit mit seiner Freundin führt. Da er die Streitereien gründlich satt hat, möchte er die Entscheidung, ob er seine Freundin weiterhin beim Shoppen begleitet, vom Ergebnis eines statistischen Tests ($\alpha = 1\ \%$) abhängig machen. Wenn die durchschnittliche Wartezeit tatsächlich 8 Minuten oder länger dauert, will er Statistica in Zukunft nicht mehr begleiten. Dabei möchte er das Risiko, sich fälschlicherweise für eine weitere Begleitung zu entscheiden, so klein wie möglich halten, da die dann anhaltenden Streitereien die Beziehung zwischen beiden gefährden würden.

Er geht davon aus, dass die Wartezeit eine normalverteilte Größe ist mit einer Standardabweichung von 2 Minuten.

Als Datenbasis zur Testdurchführung sollen die nächsten 16 Shoppingtouren (einfache Stichprobe) dienen.

1) Wie lautet das durchzuführende (parametrische) Testverfahren ?
2) Wie lauten die Hypothesen für diesen Test ? Begründen Sie Ihre Wahl – falls nötig – mit Hilfe einer Risikoüberlegung !
3) Wie lautet die Prüfgröße für diesen Test formal und verbal, und wie ist diese unter H_0 verteilt ?
4) Wie lautet die Testfunktion für diesen Test, und wie ist diese unter H_0 verteilt ?
5) Bestimmen Sie den Annahme- und Verwerfungsbereich dieses Tests !
6) Berechnen Sie die Wahrscheinlichkeit für eine Fehlentscheidung bei diesem Test, falls die wahre durchschnittliche Wartezeit
   a) 6 Minuten und 18 Sekunden beträgt.
   b) 8 Minuten und 6 Sekunden beträgt.
7) Bestimmen Sie den Wert der Gütefunktion bei diesem Test, falls die wahre durchschnittliche Wartezeit
   a) 6 Minuten und 50,1 Sekunden beträgt.
   b) 7 Minuten und 30 Sekunden beträgt.
8) Wie groß müsste der Stichprobenumfang mindestens gewählt werden, damit – wenn die durchschnittliche Wartezeit tatsächlich 7,5 Minuten beträgt – die Wahrscheinlichkeit für eine richtige Entscheidung den Wert 88,1 % hat ?

Nach Ablauf der 16 Shoppingtouren stellt Statistix fest, dass er insgesamt 111 Minuten an den Einkaufskassen warten musste.

9) Wie lautet die Testentscheidung ?
10) Interpretieren Sie das Testergebnis inhaltlich und statistisch exakt !
    Wird Statistix demnach weiterhin seine Freundin beim Shoppen begleiten ?
11) Welcher Fehler könnte bei der Testentscheidung unterlaufen sein ?
12) Wie hoch ist die Wahrscheinlichkeit, dass diese Testentscheidung falsch ist ?
13) Würde sich die Testentscheidung ändern, wenn Statistix (unter sonst gleichen Bedingungen) beim Testaufbau ein Signifikanzniveau von 10 % gewählt hätte ?
    Kurze Begründung !
14) Erläutern Sie kurz den Begriff der „ Trennschärfe “ eines Tests !

Aufgabe 2

Zur Zulassung für ein Seminar muss ein Eignungstest erfolgreich abgeschlossen werden. In folgender Tabelle sind die im Eignungstest erreichten Punktezahlen der Bewerber (einfache Stichprobe) notiert :

Punktezahlen von ... bis unter	Häufigkeit	Wahrscheinlichkeit ⇨ gerundet
< 52	8	0,022750 ⇨ 0,023
52 – 61	21	
61 – 70	92	
70 – 79	91	0,341345 ⇨ 0,341
79 – 88	35	
≥ 88	9	0,022750 ⇨ 0,023

Achtung : Runden Sie im weiteren Rechenverlauf stets auf drei Nachkommastellen !

A

Es soll mit Hilfe eines geeigneten statistischen Testverfahrens ($\alpha = 5\ \%$) überprüft werden, ob die Häufigkeitsverteilung für die erreichten Punktezahlen der Bewerber einer Normalverteilung mit $\mu = 70$ und $\sigma^2 = 81$ folgt.

1) Wie lautet das hier zu verwendende Testverfahren ?
2) Formulieren Sie **verbal** die Hypothesen für diesen Test ? Begründen Sie Ihre Wahl – falls nötig – mit Hilfe einer Risikoüberlegung !
3) Wie lautet die Testfunktion, und wie ist diese unter H_0 verteilt ?
4) Bestimmen Sie den Annahme- und Verwerfungsbereich dieses Tests !
5) Wie lautet die Testentscheidung ?
6) Interpretieren Sie das Testergebnis inhaltlich und statistisch exakt !

Angenommen, man hätte überprüfen wollen, ob die Häufigkeitsverteilung für die erreichten Punktezahlen der Bewerber einer Normalverteilung folgt, wobei μ und σ^2 auf Basis des vorliegenden Stichprobenergebnisses geschätzt wurden mit $\hat{\mu} = 70$ und $\hat{\sigma}^2 = 81$.

7) Ändert sich jetzt die Testentscheidung ? Kurze Begründung !

B

Weiterhin soll mit Hilfe eines geeigneten statistischen Testverfahrens ($\alpha = 5\ \%$) überprüft werden, ob es einen geschlechtsspezifischen Unterschied hinsichtlich der erbrachten Prüfungsleistungen für das Seminar gibt.
Angenommen, folgende Informationen seien bekannt :

- insgesamt $81,25\ \%$ der Bewerber bestehen den Test.
- $68,75\ \%$ der Bewerber sind männlich.
- $80\ \%$ der weiblichen Bewerber bestehen den Test.

	bestanden	nicht bestanden	
weiblich			
männlich			

1) Wie lautet das hier zu verwendende Testverfahren ?
2) Formulieren Sie **verbal** die Hypothesen für diesen Test ? Begründen Sie Ihre Wahl – falls nötig – mit Hilfe einer Risikoüberlegung !
3) Wie lautet die Testfunktion, und wie ist diese unter H_0 verteilt ?
4) Bestimmen Sie den Annahme- und Verwerfungsbereich dieses Tests !
5) Wie lautet die Testentscheidung ?
6) Interpretieren Sie das Testergebnis inhaltlich und statistisch exakt !

Aufgabe 3

Der Kneipenbesitzer Willy hat einen neuen Spielautomaten. Pro Spielrunde muss man 1 € als Spieleinsatz einwerfen. Wird das Spiel verloren, verbleibt der Einsatz im Automaten. Beim Unentschieden erhält man seinen Spieleinsatz zurück. Gewinnt man das Spiel, bekommt man zusätzlich zum rückerstatteten Spieleinsatz eine Gewinnprämie in Höhe von 1 € ausbezahlt.
Willy interessiert sich nun für die Zufallsvariable

X : „ Gewinn pro Spielrunde [in €] "

Leider ist über die Wahrscheinlichkeiten von Verlust (π_V), Unentschieden (π_U) und Gewinn (π_G) nichts weiter bekannt. Folgende drei Zufallsvariablen seien definiert :

V : „ Anzahl der verlorenen Spiele bei n unabhängigen Spielrunden "

U : „ Anzahl der unentschiedenen Spiele bei n unabhängigen Spielrunden "

G : „ Anzahl der gewonnenen Spiele bei n unabhängigen Spielrunden "

x	− 1	0	+ 1
P (X = x)	π_V	π_U	π_G

1) Bestimmen Sie den Erwartungswert der Zufallsvariablen V !
2) Stellen Sie (in Abhängigkeit von v) die Kleinste–Quadrate–Funktion für die Verlustwahrscheinlichkeit π_V auf !
3) Bestimmen Sie den Schätzwert für π_V nach der Methode der kleinsten Quadrate !
4) Wie lautet demnach die Kleinste–Quadrate–Schätzfunktion $\hat{\Theta}_{KQ}$ für π_V ?

Willy hat sich folgende Schätzfunktion $\hat{\Theta}_W$ für die Verlustwahrscheinlichkeit π_V überlegt :

$$\hat{\Theta}_W = \frac{1}{2 \cdot n} \cdot \sum_{i=1}^{n} \left(X_i^2 - X_i \right)$$

5) Bestimmen Sie den Erwartungswert der Zufallsvariablen X !
6) Geben Sie tabellarisch die Wahrscheinlichkeitsverteilung der Zufallsvariablen X^2 an !
7) Bestimmen Sie den Erwartungswert der Zufallsvariablen X^2 !
8) Prüfen Sie mit Hilfe eines formalen Rechenganges, ob $\hat{\Theta}_W$ eine erwartungstreue Schätzfunktion für die Verlustwahrscheinlichkeit π_V ist !

Willys Stammgast Kalle hat früher in der Firma gearbeitet, die den Spielautomaten hergestellt hat. Er kann sich zwar nicht mehr genau an die programmierte Wahrscheinlichkeitsverteilung erinnern, aber er weiß noch, dass die Wahrscheinlichkeit für ein Unentschieden doppelt so hoch wie für einen Verlust programmiert war.

9) Geben Sie tabellarisch (in Abhängigkeit von π_V) die sich aufgrund von Kalles Informationen ergebene Wahrscheinlichkeitsverteilung der Zufallsvariablen X an !
10) Ermitteln Sie den Maximum–Likelihood–Schätzwert für π_V (in Abhängigkeit von v , u und g) !
11) Wie lautet demnach die Maximum–Likelihood–Schätzfunktion für π_V ?
12) Prüfen Sie mit Hilfe eines formalen Rechenganges, ob $\hat{\Theta}_{ML}$ eine erwartungstreue Schätzfunktion für die Verlustwahrscheinlichkeit π_V ist !

Aufgabe MC

Block A

1) Die Varianz der Differenzfunktion X - Y bei verbundenen Stichproben kann größer sein als bei unverbundenen Stichproben.

2) Von zwei Schätzfunktionen $\hat{\Theta}_1$ und $\hat{\Theta}_2$ ist diejenige wirksamer zum Schätzen des unbekannten Parameters ϑ, die die kleinere Varianz besitzt.

3) Eine Zufallsvariable X, für die gilt: $E(X) = 0$ und $Var(X) = 1$, ist standardnormalverteilt.

4) Das „ Likelihood - Prinzip " ist nicht frequentistisch, d.h. die Bewertung der ϑ - Werte durch die Likelihoodfunktion richtet sich allein nach der einen beobachteten Stichprobe. Nach der Beobachtung wird nicht mehr berücksichtigt, was sonst noch alles hätte beobachtet werden können.

Block B

1) Das sog. „ empirische " Signifikanzniveau kann größer sein als das vorgegebene Signifikanzniveau eines Tests.

2) Eine Zufallsauswahl, die uneingeschränkt ist, muss keine gleichgewichtete Zufallsauswahl sein.

3) Bei der Intervallschätzung bezieht sich die Wahrscheinlichkeitsaussage in Form des Konfidenzniveaus $1 - \alpha$ stets auf die Situation vor der Beobachtung.

4) Sei X_1 verteilt nach $N(\mu_1 ; \sigma_1^2)$ und X_2 verteilt nach $N(\mu_2 ; \sigma_2^2)$. Dann ist $X_1 - X_2$ verteilt nach $N(\mu_1 - \mu_2 ; \sigma_1^2 + \sigma_2^2)$.

Block C

1) Die Maximum-Likelihood-Methode besagt, dass zu einem festen Stichprobenergebnis $(x_1, \ldots, x_n)$ derjenige Schätzwert für den unbekannten (zu schätzenden) Parameter ϑ zu wählen ist, unter dem im vorhinein die Wahrscheinlichkeit für das Eintreten dieses Stichprobenergebnisses am größten ist.

2) Ein Chi² - Unabhängigkeitstest kann nicht für unterschiedlich skalierte Merkmale durchgeführt werden.

3) Bei einem proportional geschichteten Auswahlverfahren hat jede Stichprobe $(X_1, \ldots, X_n)$ die gleiche Wahrscheinlichkeit realisiert zu werden.

4) Nach dem „ Mean Square Error - Konzept " soll der (quadrierte) zu erwartende theoretische Schätzfehler minimiert werden.

Block D

1) Bei einem zweiseitigen Test auf π ist die Gütefunktion nicht symmetrisch um π_0.

2) Der Wert der Gütefunktion eines Tests an der Stelle $\vartheta \in H_1$ ist das Komplement der Wahrscheinlichkeit für den Fehler 2. Art an der Stelle ϑ.

3) Das Schätzintervall für den unbekannten Parameter μ einer normalverteilten Grundgesamtheit wird bei bekanntem σ^2 (unter sonst gleichen Bedingungen) mit zunehmendem Stichprobenumfang stets kleiner.

4) Das Verfahren der Intervallschätzung liefert Schätzintervalle, in denen der gesuchte Parameter mit $(1 - \alpha)$ %-iger Wahrscheinlichkeit liegt.

Block E

1) Das (schwache) Gesetz der großen Zahlen besagt, dass die Wahrscheinlichkeit, mit der $\overline{X}_n$ in ein (beliebig klein) vorgegebenes Intervall $[\mu - c ; \mu + c]$ fällt, mit wachsender Anzahl der Versuche gegen Null konvergiert.

2) Bei einem Einstichproben - Gaußtest auf den Parameter μ ist die Wahrscheinlichkeit für den Fehler 1.Art am größten, falls gilt: $\mu = \mu_0$.

3) Fällt bei einem Parametertest der Wert der Testfunktion in den Annahmebereich, kann ein Fehler 1.Art begangen worden sein.

4) An der Gütefunktion lässt sich in Abhängigkeit vom zu testenden Parameter ablesen, mit welcher Wahrscheinlichkeit eine Nullhypothese beibehalten wird.