**Übungsblatt 3: Hypothesentests**

**Aufgabe 1:**

Bei statistischen Hypothesentests spricht man von der Wahrscheinlichkeit einen Fehler 1. Art bzw. der Wahrscheinlichkeit einen Fehler 2. Art zu machen und der Macht eines Tests. Was versteht man unter den Begriffen? Stellen Sie diese graphisch dar. Erläutern Sie die den unvermeidlichen Trade-Off zwischen den beiden Fehlerwahrscheinlichkeiten. (siehe SCHIRA, Seite 497)

**Aufgabe 2:**

In den Fällen (a)-(d) sind verschiedene Teststatistiken dargestellt. Wählen Sie aus den möglichen Interpetationen 1-8 auf der nächsten Seite für jeden Fall jeweils eine aus. Ein Satz kann für mehrere Antworten verwendet werden, einige Sätze sind unsinnig.

**(a)** Unter der Nullhypothese ist eine Teststatistik standardnormalverteilt. Die Formulierung der Alternativhypothese impliziert einen zweiseitigen Test (d.h. Ablehnung der Nullhypothese für große und kleine Werte der Teststatistik), das Signifikanzniveau ist 0,05. Die Berechnung der Teststatistik liefert den Wert -1,2. (Hinweis: Beachten Sie die Symmetrieeigenschaft der Dichtefunktion der Normalverteilung, d.h. $1 - F_X(x) = F_X(-x)$), wobei $F_X(x)$ die Verteilungsfunktion der Normalverteilung bezeichnet.

**(b)** Unter der Nullhypothese ist eine Teststatistik $\chi^2(1)$-verteilt ($\chi^2$ mit einem Freiheitsgrad). Das Signifikanzniveau ist 0,01. Die Alternativhypothese impliziert einen einseitigen Test (d.h. Ablehnung von $H_0$ für große Werte der Teststatistik). Eine Berechnung der Teststatistik liefert den Wert 7,1.

**(c)** Testdesign wie in (b): Die Berechnung der Teststatistik liefert ein empirisches Signifikanzniveau ($p$-Wert) von 0,015.

**(d)** Unter der Nullhypothese ist eine Teststatistik Student-$t$-verteilt mit 4 Freiheitsgraden. Die Alternativhypothese impliziert einen einseitigen Test, das Signifikanzniveau ist 0,01. Die Berechnung der Teststatistik liefert einen Wert von 1,9.

**Mögliche Interpretationen**

1. Die Nullhypothese kann auf dem 5% Signifikanzniveau nicht verworfen werden.
2. Die Nullhypothese ist mit einer Wahrscheinlichkeit von 95% richtig und wird daher angenommen.
3. Die Nullhypothese kann auf dem 1% Signifikanzniveau nicht verworfen werden.
4. Die Nullypothese ist wahr und wird daher angenommen.
5. Die Nullhypothese wird auf dem 5% Signifikanzniveau verworfen.
6. Die Nullhypothese ist mit einer Wahrscheinlichkeit von nur 5% richtig und wird daher verworfen.
7. Die Nullhypothese ist falsch und wird daher verworfen.
8. Die Nullhypothese wird auf dem 1% Signifikanzniveau verworfen.

**Aufgabe 3:**

Interpretieren Sie folgende Testergebnisse.

**(a)** Mit Hilfe des sogenannten Jarque-Bera Tests kann man die Nullhypothese testen, dass die Zufallsvariable nicht normalverteilt ist. Zur Konstruktion der Teststatistik wird eine Zufallsstichprobe aus der Grundgesamtheit benötigt. Unter der Nullhypothese ist die Jarque-Bera Statistik $\chi^2$-verteilt mit 2 Freiheitsgraden. Der Test ist einseitig, d.h. man verwirft die Nullhypothese für große Werte der Teststatistik. Die Berechnung der Jarque-Bera Teststatistik auf Basis einer Zufallsstichprobe von logarithmierten individuellen Einkommensdaten liefert einen Wert von 10,9. Interpretieren Sie dieses Testergebnis.

**(b)** Die Jarque-Bera Teststatistik wurde außerdem auf Basis einer weiteren Zufallsstichprobe berechnet, in welcher der IQ des Befragten erhoben wurde. Für diese Daten ergibt sich ein $p$-Wert (empirisches Signifikanzniveau) der Jarque-Bera Statistik von 0,12. Interpretieren Sie dieses Testergebnis.

**(c)** Der Kruskal-Wallis-Test wird dazu verwendet, die Nullhypothese zu testen, dass zwei oder mehr Zufallsstichproben aus Verteilungen gezogen wurden, in der die interessierende Zufallsvariable den gleichen Erwartungswert besitzt. So könnte man z.B. testen, ob die erwartete Präferenz für ein Produkt (z.B. Weißbier) in zwei Verkaufsregionen (z.B. Schleswig-Holstein und Bayern) identisch ist. Zur Konstruktion benötigt man Zufallsstichproben aus den beiden Verkaufsregionen. Unter der Nullhypothese folgt die Teststatistik einer Verteilung, die in vielen Lehrbüchern tabeliert ist. Der Test ist als einseitiger Test formuliert, d.h. wir verwerfen für große Werte der Teststatistik. In einer Marketingstudie wurden in Zufallsstichproben Daten zu den Präferenzwerten für das Produkt in den o.g. Regionen erhoben und die Kruskal-Wallis-Teststatistik berechnet. Es ergab sich ein $p$-Wert (empirisches Signifikanzniveau) von 0,001. Interpretieren Sie dieses Test-Ergebnis.

**(d)** Lo und MacKinlay haben einen Test entwickelt, mit dem die Null-Hypothese getestet werden kann, dass Aktienrenditen mit Hilfe der Historie der Renditen nicht prognostizierbar sind. Die Alternativhypothese lautet, dass Prognostizierbarkeit vorliegt. Unter der Nullhypothese ist die Lo/MacKinlay-Teststatistik standardnormalverteilt. Zur Konstruktion der Teststatistik benötigt man eine Zeitreihe von Aktienrenditen. Der Test ist zweiseitig. Unter Verwendung von Kurszeitreihen der Börse in New York wurde für eine Zeitreihe von Aktienrenditen ein Wert der Lo/MacKinlay-Teststatistik von -1,56 ermittelt. Interpretieren Sie dieses Testergebnis.

**(e)** Mit Hilfe des sogenannten KK-Test kann man die Nullhypothese testen, dass eine Kovarianz zweier Zufallsvariablen gleich Null ist. Die Alternativhypothese lautet, dass die Kovarianz ungleich Null ist. Der Test ist zweiseitig. Unter der Nullhypothese ist die Teststatistik $t$-verteilt mit 4 Freiheitsgraden. In einer konkreten Anwendung berechnet sich der Wert der der Teststatistik mit 5,3. Ermitteln Sie die kritischen Werte bei einem Signifikanzniveau von $\alpha$=0,05. Interpretieren Sie das Testergebnis.

**(f)** Eine weitere Anwendung (andere Stichprobe) des Tests aus (a) liefert einen $p$-Wert von 0,0013. Interpretieren Sie das Testergebnis.

**(g)** Eine weitere Anwendung (andere Stichprobe) des Tests aus (a) liefert einen $p$-Wert von 0,27. Interpretieren Sie das Testergebnis.

**(h)** Mit Hilfe eines "Likelihood-Ratio"-Tests können alternative Verteilungsannahmen gegeneinander getestet werden. Dabei wird eine Verteilungsannahme als restriktive Version einer anderen formuliert. So ist z.B. $X \sim N(0, \sigma^2)$ eine restriktive Version von $X \sim N(\mu, \sigma^2)$. Die Nullhypothese beim "Likelihood-Ratio"-Test lautet, dass die restriktive Verteilngsannahme richtig ist (im Beispiel, dass $\mu = 0$). Die Alternativhypothese lautet, dass diese Restriktionen falsch sind (im Beispiel, dass $\mu \neq 0$). Unter der Nullhypothese ist dann die Prüfgröße, die "Likelihood-Ratio-Statistik", $\chi^2$-verteilt mit der Anzahl an Freiheitsgraden gleich der Anzahl der Restriktionen (im Beispiel also 1). Der "Likelihood-Ratio"-Test ist einseitig, d.h. man verwirft die Nullhypothese für große Werte der Teststatistik. In einer konkreten Anwendung mit einer Parameterrestriktion errechnet sich der Wert der "Likelihood-Ratio"-Testsstatistik mit 2,051. Berechnen Sie den $p$-Wert. Was sagt der $p$-Wert aus? Erläutern Sie. Interpretieren Sie das Testergebnis unter Annahme eines von Ihnen gewählten Signifikanzniveaus.

**(i)** Ein Wissenschaftler führt zwei (zweiseitige) $t$-Tests auf Signifikanz eines geschätzten Parameters $\hat{\theta}$ durch. Der Parameter wurde mit der ML Methode geschätzt. Die zu testende Nullhypothese ist dabei einmal, dass der Grundgesamtheitsparameter $\theta$ gleich Null ist, bzw. einmal, dass $\theta = 1$. Der Stichprobenumfang ist groß, $n > 1000$. Für den ersten Test ($H_0 : \theta = 0$) ergibt sich ein Wert der $t$-Statistik von 12,3. Für den zweiten Test (der $H_0 : \theta = 1$) ergibt sich ein Wert der $t$-Statistik von -0,45. Interpretieren Sie das Testergebnis unter der Anwendung der Faustregel $|t| > 2$.

**(j)** In der Marketing-Abteilung eines Unternehmens soll überprüft werden, ob eine Erneuerung einer Produktverpackung den Verkauf des Produktes steigert. Dazu werden zwei Personengruppen zu Ihrer Kaufbereitschaft befragt. Die erste Gruppe besteht aus den 13-19 Jährigen und die zweite aus 20-30 Jährigen. Aus der Untersuchung beider Gruppen berechnet die Marketing-Abteilung zwei Teststatistiken. Unter der Nullhypothese, dass die neue Verpackung keine verkaufssteigernde Wirkung hat, sind diese Prüfgrößen $t$-verteilt mit 4 Freiheitsgraden. Der Test ist zweiseitig. Für die erste Probandengruppe (Altersgruppe 13-19) errechnet sich ein Wert für die Prüfgröße von 4,6. Für die zweite Probandengruppe (Altersgruppe 20-30) errechnet sich ein Wert für die Prüfgröße von 1,53. Berechnen Sie die $p$-Werte und interpretieren Sie die Test-Ergebnisse. Was sagt der $p$-Wert aus?