Klausur Stochastik und Statistik (WS 2007/08)

Name: Matrikelnummer:

Fachrichtung:

Unter der folgenden Nummer finden Sie Ihr Ergebnis später im Internet: | 8 | A | | | |

Aufgabe 1: Aufgabe: *Datenanalyse und Tests*

Der folgende Datensatz beschäftigt sich mit der Optimierung der Produktionsprozesse eines Bekleidungsherstellers. Eine Arbeitsgruppe hat ein computerbasiertes Verfahren entwickelt, mit dessen Hilfe der Verschnitt (also die nicht genutzte Menge an Stoff) minimiert werden soll. Bisher wurde diese Aufgabe der optimalen Stoffaufteilung von den Arbeitern durchgeführt. Die Arbeitsgruppe möchte nachweisen, dass ihr neues Verfahren im Mittel weniger Verschnitt produziert.

Dazu wurden Zuschnittsituationen aus der Menge aller (im Beobachtungszeitraum) anfallenden Zuschnittsituationen zufällig ausgewählt. Für jede dieser Zuschnittsituationen erarbeitete jeweils der zuständige Zuschneider und der Computer einen Schnittplan. Diese beiden Schnittpläne wurden verglichen und ermittelt wieviel Verschnitt durch den Computerplan eingespart worden wäre.

Der Datensatz enthält zwei Variablen:

- `Verbesserung` ist die Differenz aus der Verschnittmenge des Zuschneiders minus der Verschnittmenge des Computers. Negative Werte bedeuten, dass der Arbeiter besser war als der Computer.
- `Zuschneider` gibt an, welcher Arbeiter den Zuschnitt durchgeführt hat.

```
> Verschnitt <- read.table("textil.txt", header = TRUE)
> Verschnitt

   Verbesserung Zuschneider
1           1.3           1
2           3.4           3
3           4.6           3
4           2.4           1
5           9.2           5
6           4.2           2
7           0.3           1
8           1.2           1
```

```
9           12.3          5
10           8.5          2
11           2.9          2
12           2.8          1
13          17.0          2
14          -0.8          3
 [ erreichte getOption("max.print") -- letzte 81 Zeilen ausgelassen ]]
```

(1) Geben Sie das Skalenniveau der Variable `Verbesserung` an!

---

(2) Geben Sie das Skalenniveau der `Zuschneider` an!

---

(3) Geben Sie die Grundgesamtheit an, für die dieser Datensatz repräsentativ ist!

```
> par(mfrow = c(1, 2))
> boxplot(Verbesserung ~ Zuschneider, data = Verschnitt, xlab = "Zuschneider",
+     ylab = "Verbesserung")
> qqnorm(Verschnitt$Verbesserung)
```

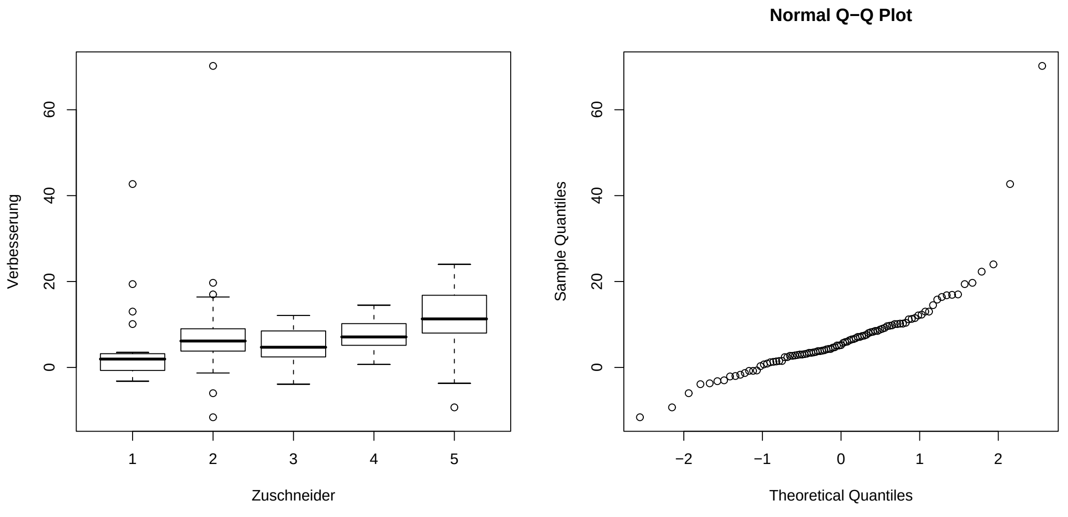

Abbildung 1: Abhängigkeit von Zuschneider und Verschnitt und der QQ-Plot aller Verschnittdaten.

(4) Welche der folgenden statistischen Graphiken eignen sich *besonders gut* zum Vergleich, wie oft jeder der 5 Zuschneider den Zuschnitt vorgenommen hat?

- **O** Streudiagramm
- **O** Balkendiagramm
- **O** einfaches Punktdiagramm
- **O** gesplittetes gestapeltes Punktdiagramm
- **O** Tortendiagramm
- **O** gesplitteter Boxplot
- **O** QQ-Plot
- **O** einfacher Boxplot
- **O** Histogramm

(5) Wie heißt die linke Graphik in Abbildung 1? (2)

---

(6) Welche der folgenden Vermutungen würden Sie aus Abbildung 1 ableiten?

- **O** Die Verteilungen sind stark assymetrisch.
- **O** Die Verteilungen haben eine stark unterschiedliche Streuung.
- **O** Es gibt mehrere Ausreißer.
- **O** Es gibt genau einen Ausreißer.
- **O** Der Zuschneider 1 scheint, im Vergleich mit dem Computer, besser zu sein als die anderen Zuscheider.
- **O** Es kann aufgrund dieser Graphik *statistisch nachgewiesen* werden, dass die Mediane der 5 Grundgesamtheiten unterschiedlich sind.
- **O** Wegen der ungleichen Mediane ist die Anwendung von normalverteilungsbasierten Verfahren anzuraten.
- **O** Der Datensatz ist – abgesehen von eventuellen Ausreißern – ungefähr normalverteilt.
- **O** Im Datensatz treten massiv viele Bindungen auf.

(7) Welcher Test würde sich eignen, um nachzuweisen dass die Zuschneider unterschiedlich effizient arbeiten?

- **O** Gumble-Watson-Test
- **O** Kruskal-Wallis-Test
- **O** Shapiro-Wilk-Test
- **O** Ansari-Friedmann-Test
- **O** χ^2-Unabhängigkeitstest
- **O** Varianzanalyse

- O Zwei Stichproben t-Test
- O Wilcoxen-Vorzeichen-Rang Test
- O F-Test

(8) Für das Testproblem zum Vergleich der Verschnitte der Menschen und des Computers würde sich in dieser Situation mit gepaarten Stichproben ein Wilcoxen-Vorzeichen-Rang Test eignen. Da die Differenzen innerhalb der Paare bereits gebildet sind, verwenden wir dazu den folgenden Befehl:

```
> wilcox.test(Verschnitt$Verbesserung, alternative = "greater")

        Wilcoxon signed rank test with continuity correction

data:  Verschnitt$Verbesserung
V = 4186.5, p-value = 7.476e-13
alternative hypothesis: true location is greater than 0
```

Welche Aussagen sind richtig? ($\alpha = 0.05$)

- O Der Test hat die Hypothese signifikant abgelehnt.
- O Der Test hat die Alternative signifikant abgelehnt.
- O Der Test hat die Hypothese angenommen.
- O Der Test hat die Hypothese signifikant angenommen.
- O Der Test setzt eine repräsentative Stichprobennahme voraus.
- O Der Test setzt die Normalverteilung der Daten voraus.
- O Der Test hat Probleme, wenn viele Bindungen vorliegen.
- O Die Voraussetzungen des Tests sind erfüllt.
- O Die Voraussetzungen des Tests sind nicht erfüllt.

(9) Formulieren Sie das Endergebnis der Aufgabe in Bezug auf die ursprüngliche Fragestellung in allgemeinverständlichen Begriffen!

Aufgabe 2: *Modellieren*
Ordnen Sie den folgenden Verteilungen und Methoden ihren Anwendungssituationen zu:

a: Binomialverteilung

b: Hypergeometrische Verteilung

c: Faltung

d: Normalverteilung

e: Poisson-Verteilung

f: $F_1(t)F_2(t)F_3(t)$

g: Fréchet-Verteilung

☐ Anzahl der Glühbirnenausfälle bei einer großen Leuchtreklametafel mit mehreren tausend Glühbirnen.

☐ Die Zeit die ein Bagger benötigt, um 100 Wagons zu beladen.

☐ Verteilung der Reperaturdauer, wenn vor dem Start des Flugzeugs drei Dinge nacheinander repariert werden müssen und die Reperaturzeitverteilungen für die einzelnen Schritte gegeben ist.

☐ Verteilung der Reperaturdauer, wenn in der gleichen Situation die drei Dinge gleichzeitig repariert werden können.

☐ Die maximale Sturmstärke, die der Eifelturm in den nächsten 50 Jahren aushalten muss.

☐ Im Rahmen der Qualitätskontrolle wird jeden Tag ein Test mit dem Niveau $\alpha = 5\%$ durchgeführt, wobei die Hypothese einem funktionierenden Produktionsablauf entspricht. Gesucht ist die Verteilung der Anzahl der in einer Woche (=5Tage) abgelehnten Hypothesen, wenn die Produktion tatsächlich funktioniert.

Aufgabe 3: *Zuverlässigkeit einer Traglufthalle*
Eine Traglufthalle ist ständig auf Energieversorgung angewiesen, um den Innendruck aufrecht zu erhalten, der die Halle stabil hält. Wird der Innendruck nicht aufrecht erhalten bricht die Halle zusammen und es entsteht großer Schaden.

In der Halle findet über einen Zeitraum von 14 Tagen eine Dinosaurierausstellung statt, zu der viele Besucher erwartet werden. Der Veranstalter möchte aus Angst um die Besucher und seine teuren Ausstellungstücke eine Zuverlässigkeitsrechung für den Ausfall der Energieversorgung haben. Die Stadtwerke geben die Wahrscheinlichkeit für einen Ausfall der Vorsorgung über das Stromnetz im Veranstaltungszeitraum mit ≤ 0.01 an. Fällt die Versorgung über das Netz aus, so wird das System der Traglufthalle automatisch zwei Notstromaggregate zu starten versuchen. Erfahrungsgemäß starten die Motoren der Aggregate unabhängig voneinander mit einer Wahrscheinlichkeit von ≥ 0.95. Ein gestartetes Aggregat kann die Versorgung der Halle übernehmen. (Der Fall mehrfacher Stromausfälle oder der spätere Ausfall eines Aggregats braucht nicht betrachtet zu werden, da dies sehr unwahrscheinlich ist.)

(1) Wie hoch ist die Wahrscheinlichkeit, dass die Stromversorgung der Halle im Veranstaltungszeitraum zusammenbricht, höchstens?

---

(2) Die Wahrscheinlichkeit für den Ausfall des aus mehreren Lüftern bestehenden Belüftungssystem (also der Lüfter, welche die Energie verbrauchen, um die Luft in die Halle zu blasen) wird mit 0.00001 angegeben. Berechnen Sie die Wahrscheinlichkeit, dass die Halle während der gesamten Veranstaltung mit Luft versorgt wird!

Aufgabe 4: *Grundlagen*

(1) Die Lebensdauerverteilung eines Bauteils in Jahren ist durch $F_X(x) = 1 - 2^{-x}$ gegeben. Wie groß ist die Wahrscheinlichkeit, dass dieses Bauteil 3 Jahre durchhält?

---

(2) Wieviele Möglichkeiten gibt es, ...

- eine Gruppe mit n Frauen und n Männer in n Tanzpaare aufzuteilen.

- einen Sitzplan für n Gäste aufzustellen, wenn es $m > n$ Plätze gibt.

- ein Gerät einzustellen, das 5 Kippschalter mit jeweils zwei Einstellmöglichkeiten hat.

- Wieviele Möglichkeiten gibt es 3 von 7 Monteuren für einen Einsatz auszusuchen?

---

(3) Wie wahrscheinlich ist es ...

- einen 6-er Pasch zu würfeln (d.h. man würfelt mit zwei Würfeln und beide Würfeln zeigen eine 6)?

- eine Prüfung zu bestehen, wenn es zwei Prüfungsgebiete gibt, die jeweils mit der Wahrscheinlichkeit $\frac{1}{2}$ ausgewählt werden und man das erste so gut kann, dass man mit einer Wahrscheinlichkeit von 70% durchkommt und beim anderen nur mit Wahrscheinlichkeit von 20%?

Aufgabe 5: Aufgabe: Firma Knauser
Sie sind bei der Firma Knauser und Co. KG angestellt und sollen entscheiden, ob die produzierten Koffer mit billigen Plastikrollen (Preis: 1Euro) oder teuren Aluminumrollen (Preis: 3Euro) ausgestattet werden sollen. Die Aluminumrollen halten mit der Wahrscheinlichkeit von 0.99 und die Plastikrollen mit einer Wahrscheinlichkeit von 0.8. Gehen die Rollen kaputt, muss der Kaufpreis von $50 Euro$ zurückerstattet werden.

(1) Welche Rollen führt bei ihrem Einbau zu einem besseren erwarteten Betriebsergebnis? (Wenn man davon absieht, dass die Kunden möglicherweise nie wieder Knauser kaufen, wenn die Rollen kaputt gegegangen sind. Aber das täten sie sonst ja auch nicht, da sie dann ihren Koffer ja noch haben.)

Zu dieser Aufgabe wird eine ausführliche Nebenrechnung und die Berechnung von Erwartungswerten verlangt. Falls der Platz nicht reicht, geben Sie bitte einen mit Namen, Studiengang und Matrikelnummer markierten Zusatzzettel ab.