

## MODUL E - Übungsaufgaben

### Aufgabe E.1

Um die Qualität des Betons auf einer Baustelle zu testen, wird die Druckfestigkeit des hergestellten Betons getestet. Erfahrungsgemäss ist bekannt, dass die Druckfestigkeit einer Normalverteilung folgt und die Varianz der Druckfestigkeit für diese Betonsorte  $16.36 \text{ [MPa}^2\text{]}$  ist. Akzeptanzkriterium für die Qualität des Betons auf der Baustelle ist, dass der Mittelwert der Betondruckfestigkeit gleich  $30 \pm \Delta$  (MPa) ist. Dies wird täglich am jeweils hergestellten Beton gemessen. Um die Homogenität der Verarbeitung zu gewährleisten, sind sowohl kleinere wie auch grössere Werte nicht akzeptabel.

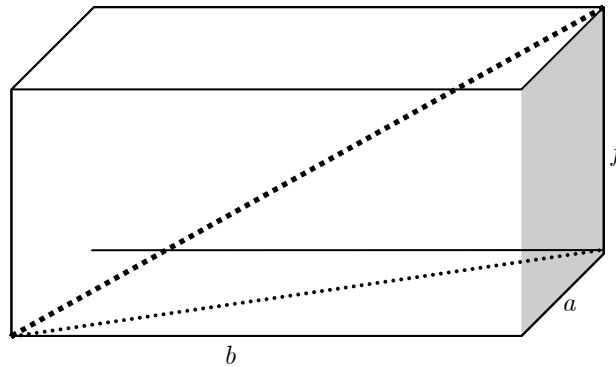
Aus einer Tagesproduktion werden 15 Proben entnommen und ihre Druckfestigkeit getestet. Die Ergebnisse sind in Tabelle E.1.1 wiedergegeben. Kann die Qualität des Betons akzeptiert werden? Teste die Hypothese jeweils für ein Signifikanzniveau von 10 % und 1 %.

**Tabelle E.1.1:** Gemessene Druckfestigkeit der 15 Proben.

Nummer der Probe ( $i$ )	Druckfestigkeit (MPa)
1	24.4
2	26.5
3	27.8
4	29.2
5	39.2
6	37.8
7	35.1
8	30.8
9	30.3
10	39.7
11	38.4
12	33.3
13	33.5
14	28.1
15	34.6

## Aufgabe E.2

Gegeben sei ein Quader, bei welchem die Kanten  $a$ ,  $b$  und  $f$  (Abb. D.13.1) gemessen wurden. Die Messungen beinhalten einen Messfehler, welcher als  $\varepsilon$  bezeichnet wird. Es wird angenommen, dass der Fehler  $\varepsilon$  zufällig (nicht systematisch) und normalverteilt ist. Die Standardabweichung des Fehlers ist  $\sigma_\varepsilon$ .



- Stelle die Wahrscheinlichkeitsdichtefunktion und die Wahrscheinlichkeitsverteilungsfunktion des Fehlers für  $d$  auf, wenn  $d$  mit Hilfe der Messungen in  $a$ ,  $b$  und  $f$  berechnet wird.
- Wenn die gleichen Messungen zur Berechnung von  $c$  verwendet werden, wie gross ist die Wahrscheinlichkeit, dass der Fehler in  $c$  grösser als  $2.4 \sigma_\varepsilon$  ist?

### Aufgabe E.3

In einem Prüflabor werden jeden Tag 30 Messungen durchgeführt, um die Qualität des Wassers zu kontrollieren. Jedes Messergebnis folgt einer Normalverteilung mit einem Mittelwert  $\mu = 23 \text{ ng / ml}$  und einer Standardabweichung  $\sigma = 4.3 \text{ ng / ml}$ .

- a) Mit welcher Wahrscheinlichkeit wird eine Messung einen Wert kleiner als  $23 \text{ ng / ml}$  aufweisen?
- b) Mit welcher Wahrscheinlichkeit liegt eine Messung im Intervall  $[19.5 \text{ ng / ml} ; 20.5 \text{ ng / ml}]$ ?
- c) Mit welcher Wahrscheinlichkeit ergibt sich ein Tagesmittelwert von kleiner als  $20 \text{ ng / ml}$ ?

### Aufgabe E.4

Ein Student liest in einem Bericht über Verkehrsanalysen, dass die mittlere Fahrzeit mit einem Personenwagen von seinem Wohnort Baden bis zur ETH Höggerberg während des Berufsverkehrs 23.7 Minuten, mit einer Standardabweichung von 3 Minuten, beträgt.

Bei seinen nächsten 13 Fahrten während des Berufsverkehrs notiert er sich seine eigenen Fahrzeiten und kommt auf einen Durchschnitt von 22.3 Minuten.

Berechne auf einem Signifikanzniveau von 5 %, ob der Bericht mit diesen Messungen übereinstimmen kann, unter der Annahme einer normalverteilten Fahrzeit und unter der Annahme, dass auch bei seinen Messungen die Standardabweichung 3 Minuten beträgt.

## Aufgabe E.5

Aus Verkehrszählungen liegt eine Datenserie vor, die den täglichen Verkehrsfluss in der Rosengartenstrasse in Zürich beschreibt (Tabelle E.5.1).

**Tabelle E.5.1:** Anzahl Fahrzeuge pro Tag, sortiert.

Tag ( $i$ )	Anzahl Fahrzeuge $\hat{x}_i$
1	3600
2	4500
3	5400
4	6500
5	7000
6	7500
7	8700
8	9000
9	9500

- a) Erstelle ein Wahrscheinlichkeitspapier für eine Dichtefunktion  $f(x) = 2/10000^2 x$  für  $0 < x < 10000$ .
- b) Überprüfe mit Hilfe des erstellten Wahrscheinlichkeitspapiers, ob der tägliche Verkehrsfluss mit dieser Verteilung angenommen werden kann.

Verwende folgende beobachtete kumulative Wahrscheinlichkeitsverteilungsfunktion:  $F_x(\hat{x}_i) = \frac{i}{N+1}$

**Aufgabe E.6**

Zur Dimensionierung eines Parkhauses wurden die Ankunftszeiten von Fahrzeugen aufgenommen. Die Zeitdifferenzen zwischen ankommenden Fahrzeugen sind in Tabelle E.6.1 gegeben.

- a) Zeichne das Wahrscheinlichkeitspapier für die Exponentialverteilung und trage die Werte der Zeitdifferenzen ein. Beurteile, ob die Intervalle einer Exponentialverteilung folgen.
- b) Bestimme den Erwartungswert der Zeitdifferenz grafisch aus der in a) erzeugten Grafik unter der Annahme, dass die Daten exponentialverteilt sind. Berechne den Erwartungswert und vergleiche die beiden Werte.

Verwende für die beobachtete kumulative Wahrscheinlichkeitsverteilungsfunktion:  $F_x(\hat{x}_i^o) = \frac{i}{N+1}$

**Tabelle E.6.1:** Zeitdifferenz zwischen den Ankunftszeiten der Fahrzeuge.

$i$	Zeitdifferenz $\hat{x}_i^o$ (Sekunden)
1	1.52
2	6.84
3	9.12
4	10.64
5	15.20
6	21.28
7	30.40
8	30.40
9	34.20
10	60.80
11	78.28
12	95.76

## Aufgabe E.7

Um die Druckfestigkeit von Beton einer bestimmten Produktionsmenge zu modellieren, wurden 20 Stichproben gemessen. Die Ergebnisse sind in der Tabelle dargestellt. Es wird angenommen, dass die Grundgesamtheit der Stichproben einer Normalverteilung  $N(\mu, \sigma)$  folgt.

**Tabelle E.7.1:** Messwerte der Druckfestigkeit von den Betonproben.

Nr. der Messung	Druckfestigkeit $\hat{x}_i^o$ [N/mm <sup>2</sup> ]
1	24.4
2	27.6
3	27.8
4	27.9
5	28.5
6	30.1
7	30.3
8	31.7
9	32.2
10	32.8
11	33.3
12	33.5
13	34.1
14	34.6
15	35.8
16	35.9
17	36.8
18	37.1
19	39.2
20	39.7

Schätze die unbekannt Parameter  $\mu$ ,  $\sigma$  anhand der Methode der Momente.

### Aufgabe E.8

Für die Stichprobe aus Aufgabe E.7 sollen nun die Parameter mit der Maximum-Likelihood-Methode (MLM) geschätzt werden.

- Beschreibe die Likelihoodfunktion.
- Schätze die Parameter  $\mu$ ,  $\sigma$  anhand der MLM.

Was passiert, wenn die Normalverteilung durch eine Exponentialverteilung ersetzt wird?

- Schätze die Parameter der Exponentialverteilung mit der MLM. Stelle die kumulative Verteilungsfunktion der Exponentialfunktion mit den geschätzten Parametern graphisch dar (z.B. Excel, Matlab), und vergleiche sie mit den beobachteten Werten.

### Aufgabe E.9

Von allen Studenten der ETH Zürich wurde das Gewicht in Kilogramm gemessen. Die Entnahme einer Stichprobe  $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)$  ergibt folgende Werte: 95, 77, 83, 71 [kg].

Das Gewicht der Studenten kann als Zufallsvariable  $X$  modelliert werden. Die Standardabweichung  $\sigma_x = 9$  [kg] sei bekannt.

- Schätze anhand der Stichprobe den Erwartungswert  $E(\bar{X})$  und die Varianz  $Var(\bar{X})$  des Stichprobenmittelwertes  $\bar{X}$ .
- Stelle die Streuung des Mittelwertes grafisch dar.
- Ermittle den Bereich, in dem der wahre Mittelwert  $\mu_x$  mit 95% Konfidenz zu erwarten ist.

### Aufgabe E.10

In einem Sägewerk wird das Schnittholz einer bestimmten Sortierklasse einer Qualitätskontrolle unterzogen. Pro Produktionstag wird eine Stichprobe mit zehn Brettern entnommen und jedes Brett auf seine Steifigkeit getestet. Der Mittelwert der Steifigkeit für jede Stichprobe ist normalverteilt und wird in Kontrolltabellen festgehalten.

Basierend auf Erfahrung kann angenommen werden, dass die Standardabweichung der Steifigkeit unabhängig von der Qualität  $\sigma = 1430 \text{ MPa}$  beträgt. Qualitätsschwankungen äussern sich nur in Form von Schwankungen des Mittelwertes.

- a) Berechne das Konfidenzintervall für die Schätzung des Stichprobenmittelwertes  $\bar{X}$  nach 15 Produktionstagen bei gegebenem  $\alpha = 0.05$ .
- b) Berechne aus a) das Konfidenzintervall für einen beobachteten Stichprobenmittelwert von  $\bar{x} = 11'000 \text{ [MPa]}$ .
- c) Wie viele Stichprobenentnahmen wären notwendig, um zu zeigen, dass der Mittelwert aller Stichproben mit mindestens  $\alpha = 0.01$  im gleichen Konfidenzintervall wie unter b) berechnet liegt?



**Aufgabe E.11**

Die Druckfestigkeit von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in Tabelle E.11.1 gegeben.

- a) Passe eine Exponentialverteilung und eine Weibullverteilung den Daten an. Bestimme dazu die Parameter dieser Verteilungen mit der Methode der Momente.
- b) Zeichne die kumulative Verteilungsfunktion der beiden Verteilungen und zeichne jeweils die kumulative Verteilung der Stichprobe ein.
- c) Teste die Güte der Anpassung für die Verteilungen mit dem  $\chi^2$ -Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in Tabelle E.11.2.
- d) Wir nehmen nun an, dass wir die Parameter für die Verteilungen aus Teilaufgabe a) der Literatur entnommen haben.  
Teste die Güte der Anpassung für die Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

**Tabelle E.11.1:** Druckfestigkeit von Holzproben.

Nr.	Druckfestigkeit [MPa]	Nr.	Druckfestigkeit [MPa]	Nr.	Druckfestigkeit [MPa]
1	12.8	11	23.4	21	29.3
2	16.3	12	26.8	22	29.5
3	16.6	13	26.9	23	30.3
4	16.9	14	27	24	32.1
5	17.2	15	27.1	25	32.3
6	17.9	16	27.2	26	33.5
7	19.5	17	27.2	27	33.9
8	21.9	18	27.5	28	35.6
9	22.3	19	27.9	29	39.2
10	22.5	20	28.3	30	43.5

**Tabelle E.11.2:** Intervalle für den  $\chi^2$ -Test

Intervall	Häufigkeit	Wahrscheinlichkeit P[Stichprobe in diesem Intervall]	erwartete Häufigkeit
-20			
20-25			
25-30			
30-			

**Hinweis:**

Die kumulative Verteilungsfunktion der Weibullverteilung kann wie folgt dargestellt werden:

$$F_x(x) = 1 - \exp\left\{-\left(\frac{x}{u}\right)^k\right\} \quad x > 0$$

Der Mittelwert und die Standardabweichung der Weibullverteilung sind gegeben als:

$$\mu = u\Gamma\left(1 + \frac{1}{k}\right) \quad \sigma = u\sqrt{\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)}$$

**Aufgabe E.12**

Anhand eines Teiles des in der ersten Vorlesung des FS2008 erhobenen Datensatzes (Tabelle E.12.1), welcher die Körpergrösse aller Studentinnen beinhaltet, soll Folgendes untersucht werden:

- a) Passe eine Normalverteilung den gegebenen Beobachtungen an. Verwende dazu den Mittelwert aus der Literatur und bestimme den Parameter  $\sigma$  mit der Maximum Likelihood Methode.
- b) Teste die Güte der Anpassung für die Normalverteilung mit dem  $\chi^2$ -Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle aus Tabelle E.12.2.

**Tabelle E.12.1:** Körpergrössen der Studentinnen

Nr.	Grösse [cm]	Nr.	Grösse [cm]	Nr.	Grösse [cm]	Nr.	Grösse [cm]
1	158	11	164	21	170	31	175
2	158	12	165	22	170	32	175
3	158	13	165	23	172	33	176
4	160	14	165	24	172	34	176
5	160	15	166	25	172	35	176
6	162	16	166	26	173	36	177
7	162	17	168	27	174	37	178
8	164	18	168	28	174	38	183
9	164	19	169	29	175		
10	164	20	170	30	175		

**Tabelle E.12.2:** Intervalle für den  $\chi^2$  Test

Intervall	Häufigkeit	Wahrscheinlichkeit P [Stichprobe in diesem Intervall]	erwartete Häufigkeit	Normalisierte Quadrate der Differenzen
0-160				
161-165				
166-170				
171-175				
176-				
Summe				

### Aufgabe E.13

Die Betondruckfestigkeit von Probekörpern einer bestimmten Produktion wird als normalverteilt angenommen. Aus früheren Testergebnissen sind der Mittelwert und die Standardabweichung der Druckfestigkeit bekannt. Nach Berücksichtigung der statistischen Unsicherheit ist der Mittelwert der Betondruckfestigkeit ebenfalls normalverteilt:

$$\mu_x \sim N(\mu_{\mu_x} = 35\text{MPa}, \sigma_{\mu_x} = 3\text{MPa})$$

Die Standardabweichung wird als bekannt (deterministisch) angenommen:

$$\sigma_x = 10\text{MPa}$$

Um die Verteilung des Parameters  $\mu_x$  zu aktualisieren, wurden 20 Versuchskörper auf ihre Druckfestigkeit geprüft. Die Ergebnisse sind in der folgenden Tabelle gegeben:

Nr.	Druckfestigkeit [MPa]	Nr.	Druckfestigkeit [MPa]
1	24.4	11	33.3
2	27.6	12	33.5
3	27.8	13	34.1
4	27.9	14	34.6
5	28.5	15	35.8
6	30.1	16	35.9
7	30.3	17	36.8
8	31.7	18	37.1
9	32.2	19	39.2
10	32.8	20	39.7

Bestimme die a posteriori Verteilung des Mittelwertes sowie die prädiktive Verteilung der Betondruckfestigkeit.

## Aufgabe E.14

Uns wurden Verkehrsdaten der Rosengartenstrasse in Zürich zur Auswertung übergeben (siehe Tabelle). Richtung 1 gibt die Verkehrsbelastung zum Bucheggplatz, Richtung 2 die Belastung zum Escher-Wyss-Platz an.

Datum	Richtung 1	Richtung 2
01.04.2001	32618	24609
02.04.2001	33380	29965
03.04.2001	34007	30629
04.04.2001	33888	30263
05.04.2001	35237	31405
06.04.2001	35843	31994
07.04.2001	33197	26846
08.04.2001	30035	22762
09.04.2001	32158	30366
10.04.2001	33406	29994
11.04.2001	34576	30958
12.04.2001	34013	30680
13.04.2001	24846	19735
14.04.2001	28252	21145
15.04.2001	25365	17805
16.04.2001	24862	18123
17.04.2001	32472	28117
18.04.2001	33245	28858
19.04.2001	33788	29080
20.04.2001	34076	30313

Der Zusammenhang der Verkehrsdaten für die beiden Richtungen soll mit Hilfe einer linearen Regression bestimmt werden.

- Erstelle ein Regressionsmodell für die Abhängigkeit zwischen den Daten der beiden Richtungen: Bestimme die Regressionskoeffizienten sowie die Varianz des Residualwertes und die Kovarianzmatrix der geschätzten Modellparameter. Verwende hierzu nur die ersten 10 Datenpaare.
- Das Regressionsmodell soll nun mit neuen Daten aktualisiert werden. Bestimme die a posteriori Regressionskoeffizienten und quantifiziere die Unsicherheit des aktualisierten Regressionsmodells. Verwende hierzu die zweite Hälfte des Datensatzes.