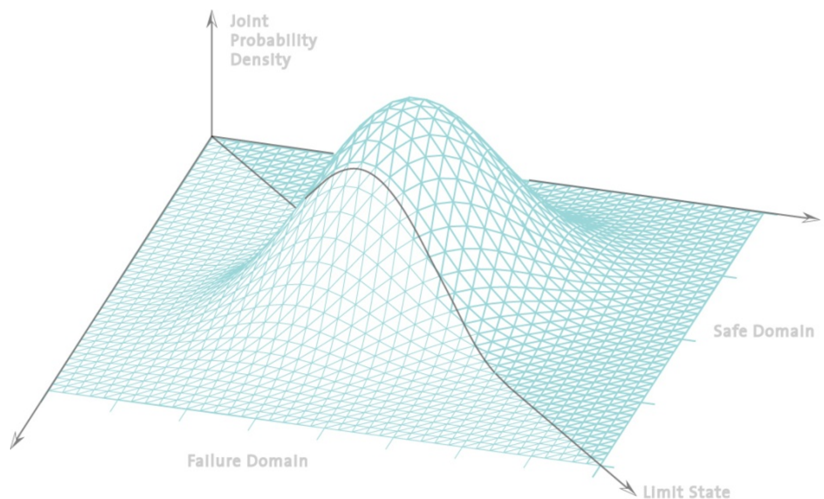


Statistik und Wahrscheinlichkeitsrechnung

Übung 10

2. Teilprüfung am Dienstag, 20. Mai

Ort und Zeit:

- Start 8:15, Ende 9:45; Dauer: 90 Minuten.
- Studierende A - Eg im HIL E 6, Eh-Ky im HIL E 9
  Studierende La - Z im HIL E 1

Erlaubte Hilfsmittel:

- Alle Unterlagen (Skripte, Bücher, andere Ausdrucke, etc.) erlaubt.
- Taschenrechner (ohne Kommunikationsmittel) erlaubt, auch programmierbare.
- Keine Kommunikationsmittel (z.B. Telefon) erlaubt.

Hinweise:

- Jede Studentin/jeder Student erhält ein Couvert mit seinem Namen, darin enthalten sind die Aufgabenblätter und ein kariertes + gestempeltes Blatt für die Lösung der Rechenaufgabe.
- Legi mitnehmen!
- Falls vor 9:15 fertig: dürft ihr euch melden und hinausgehen.
- Danach müsst ihr bis zum Prüfungsende (9:45) sitzen bleiben.

Vorlesung

Am Donnerstag findet anstatt der Übungsstunde eine Vorlesung im HIL E4 statt.

Aufgabe 9.5 Gruppenaufgabe

a) Beschreibe die Likelihood-Funktion.

b) Schätze die unbekannten Parameter (μ,σ) anhand der MLM.

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

c) Schätze die Parameter der Exponentialverteilung mit der MLM und vergleiche die kumulative Verteilungsfunktion mit den beobachteten Werten.

No. of sample	Compressive strength [MPa]	No. of sample	Compressive strength [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

Aufgabe 9.5 Gruppenaufgabe

a) Die Likelihood-Funktion ist:

$$L(\mu,\sigma \mid \hat{\mathbf{x}}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\hat{x}_i-\mu)^2}{2\sigma^2}\right]$$

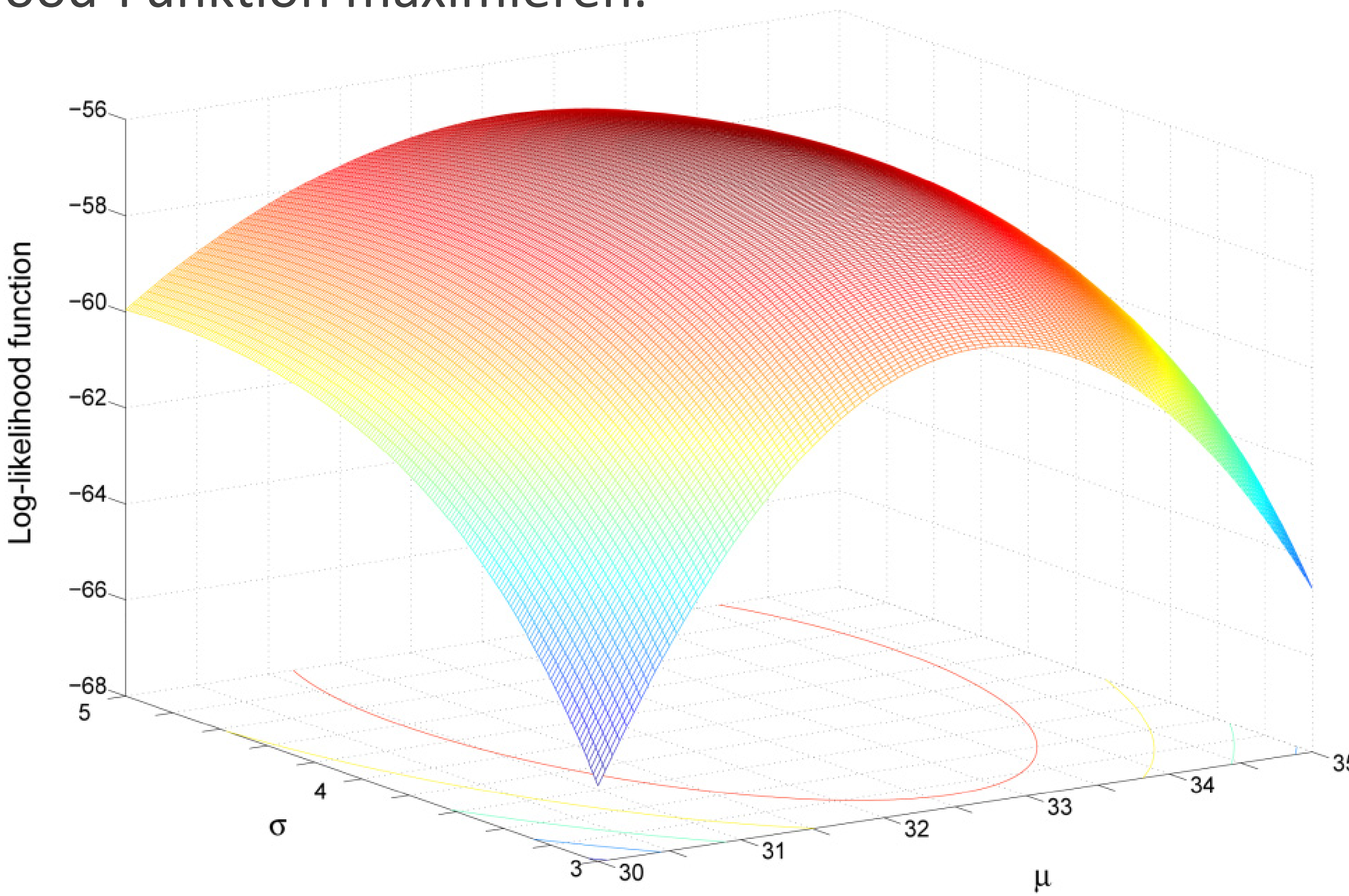
… und muss nun maximiert werden. Anstelle der Likelihood-Funktion empfiehlt es sich, die Log-Likelihood-Funktion zu verwenden:

$$l = \ln L = \sum_{i=1}^{n} \ln\left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\hat{x}_i-\mu)^2}{2\sigma^2}\right]\right]$$

$$= -n\ln(\sqrt{2\pi}) - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2$$

Aufgabe 9.5 Gruppenaufgabe

b) Die ML-Schätzwerte sind die Funktionsparameter, die die Likelihood-Funktion maximieren.



Aufgabe 9.5 Gruppenaufgabe

b) Die ML-Schätzwerte sind die Funktionsparameter, die die Likelihood-Funktion maximieren.

$$l = \ln L = \sum_{i=1}^{n} \ln\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\hat{x}_i-\mu)^2}{2\sigma^2}\right]\right]$$

$$= -n\ln(\sqrt{2\pi}) - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2$$

$$\left.\begin{aligned} \frac{\partial l}{\partial \mu} &= \frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(\hat{x}_i-\mu) = 0 \\ \frac{\partial l}{\partial \sigma} &= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2 = 0 \end{aligned}\right\} \quad \begin{aligned} \mu &= \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i \\ \sigma^2 &= \frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2 \end{aligned} \quad \rightarrow \quad \begin{aligned} \mu &= 32.67 \\ \sigma &= 4.04 \end{aligned}$$

Aufgabe 9.5 Gruppenaufgabe

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

- Schätze die unbekannten Parameter der Exponentialfunktion mit Hilfe der MLM.
- Zeichne die kumulative Verteilungsfunktion zusammen mit der beobachteten Verteilungsfunktion der Stichprobenwerte.

Aufgabe 9.5 Gruppenaufgabe

Die Likelihood-Funktion ist:

Exponentielle Verteilungsfunktion

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda \hat{x}_i}$$

Die korrespondierende Log-Likelihood-Funktion ist:

$$l = \ln L = \sum_{i=1}^{n} \ln\left(\lambda e^{-\lambda \hat{x}_i}\right) = n \ln \lambda - \lambda \sum_{i=1}^{n} \hat{x}_i$$

Den MLM Schätzwert erhält man aus:

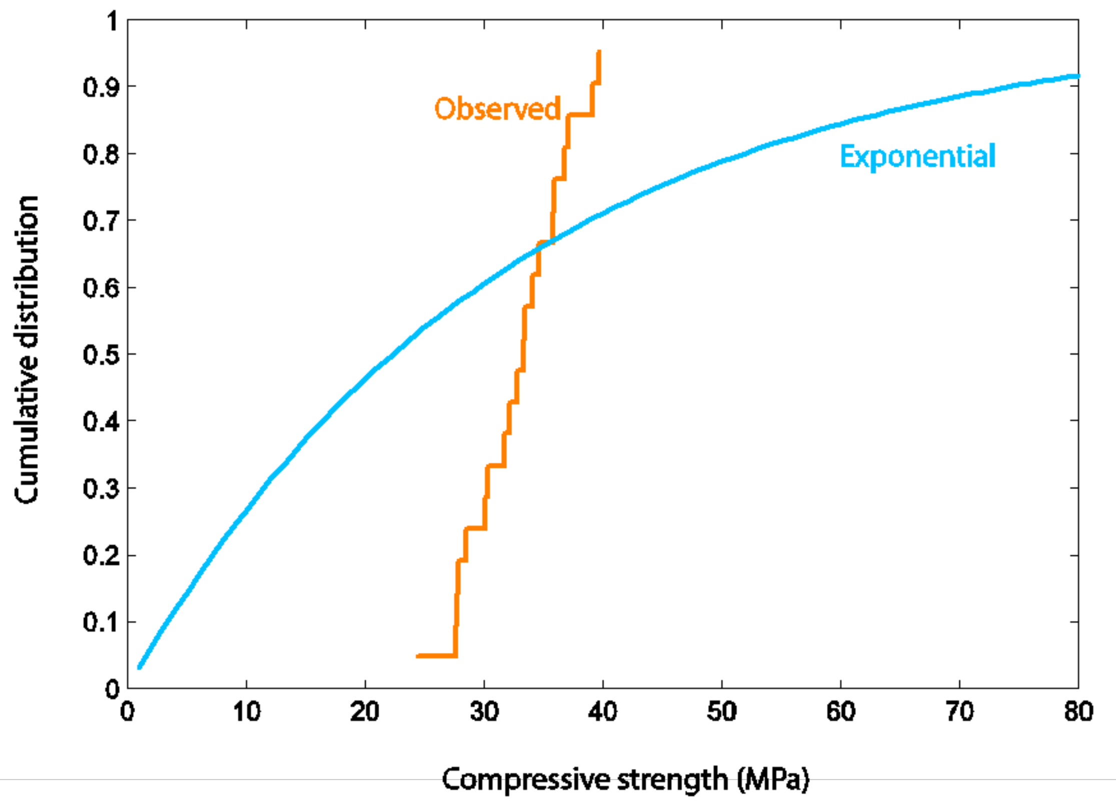
$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} \hat{x}_i = 0$$

... er entspricht:

$$\lambda = \frac{n}{\sum_{i=1}^{n} \hat{x}_i} \longrightarrow \lambda = 0.031$$

Aufgabe 9.5 Gruppenaufgabe

c) Was meint ihr? Beschreibt die Exponentialfunktion (mit den MLM Schätzwerten) die Grundgesamtheit der Stichprobe gut?



Aufgabe 10.1

Die Druckfestigkeit parallel zur Faser von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in der Tabelle gegeben.

Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [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

Aufgabe 10.1

Die Druckfestigkeit parallel zur Faser von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in der Tabelle gegeben.

a) Passe die Exponentialverteilung und die Weibullverteilung den Daten an. Bestimme dazu die Parameter dieser Verteilungen mit der Methode der Momente.

b) Zeichne die kumulativen Verteilungsfunktionen der Verteilungen und zeichne die kumulativen Verteilung der Stichprobe ein.

c) Teste die Güte der Anpassung für die Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Aufgabe 10.1

Die Druckfestigkeit parallel zur Faser von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in der Tabelle gegeben.

Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [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

a) Passen Sie die Exponentialverteilung und die Weibullverteilung den Daten an. Bestimmen Sie dazu die Parameter dieser Verteilungen mit der Methode der Momente.

Aufgabe 10.1

Das erste und zweite Moment der Beobachtungen sind:

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i = 26.41$$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i^2 = 747.55$$

Die Exponentialverteilung hat die folgende Verteilung:

$$F_X(x) = 1 - \exp(-\lambda x), \quad x > 0$$

mit dem Mittelwert:

$$\mu = \frac{1}{\lambda}$$

Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [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

Aufgabe 10.1

Das erste und zweite Moment der Beobachtungen sind:

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i = 26.41$$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i^2 = 747.55$$

Die Exponentialverteilung hat die folgende Verteilung:

$$F_X(x) = 1 - \exp(-\lambda x), \quad x > 0$$

mit dem Mittelwert:

$$\mu = \frac{1}{\lambda}$$

So kann der Parameter λ geschätzt werden zu:

$$\hat{\lambda} = \frac{1}{m_1} = 0.038$$

Die kumulative Verteilungsfunktion kann angegeben werden zu:

$$F_X(x) = 1 - \exp(-0.038x), \quad x > 0$$

Aufgabe 10.1

Das erste und zweite Moment der Beobachtungen sind:

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i = 26.41 \qquad \mu = m_1$$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i^2 = 747.55 \qquad \sigma = \sqrt{m_2 - m_1^{\;2}}$$

Die (zweiparametrische) Weibullverteilung hat die folgende kumulative Verteilungsfunktion:

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

mit dem Mittelwert und der Standardabweichung:

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

Siehe Skript Tabelle D.2

Aufgabe 10.1

Die unbekannten Parameter u und k können wie folgt bestimmt werden:

Parameter k :

$$\frac{\sigma}{\mu}=\frac{\sqrt{m_2-m_1^2}}{m_1}=\frac{u\sqrt{\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)}}{u\Gamma\left(1+\frac{1}{k}\right)}=\sqrt{\frac{\Gamma\left(1+\frac{2}{k}\right)}{\Gamma^2\left(1+\frac{1}{k}\right)}-\frac{\Gamma^2\left(1+\frac{1}{k}\right)}{\Gamma^2\left(1+\frac{1}{k}\right)}}=\sqrt{\frac{\Gamma(1+2/k)}{\Gamma^2(1+1/k)}-1}$$

$$\Gamma(p)=\int_0^\infty x^{p-1}\cdot e^{-x}dx$$

Approximativ (ohne Beweis !) gilt:

$$k\approx\left(\frac{\sigma}{\mu}\right)^{-1.09}=4.21$$

(z.B. mit Solver aus Excel)

Parameter u :

$$\mu=u\Gamma\left(1+\frac{1}{k}\right) \qquad u=\frac{m_1}{\Gamma(1+1/k)}=\frac{m_1}{\Gamma(1.24)}$$

$\Gamma(x)$	x	$\Gamma(x)$	x
.514	1.0	1.000	5.5
.863	1.1	0.951	5.6
.689	1.2	0.918	5.7
.811	1.3	0.897	5.8
.132	1.4	0.887	5.9
.591	1.5	0.886	6.0
150	1.6	0.894	6.1

Linear interpolieren

Aufgabe 10.1

Die unbekannten Parameter u und k können wie folgt bestimmt werden:

Parameter k :

$$\frac{\sigma}{\mu} = \frac{\sqrt{m_2 - m_1^{\,2}}}{m_1} = \sqrt{\frac{\Gamma\left(1+2/k\right)}{\Gamma^2\left(1+1/k\right)} - 1}$$

$$\Gamma\left(p\right) = \int_0^\infty x^{p-1} \cdot e^{-x} dx$$

Approximativ (ohne Beweis !) gilt:

$$k \approx \left(\frac{\sigma}{\mu}\right)^{-1.09} = 4.21$$

Parameter u :

$$u = \frac{m_1}{\Gamma(1+1/k)} = \frac{m_1}{\Gamma(1.24)}$$

Die geschätzten Parameter können zu:

$$\hat{k} = 4.21$$

$$\hat{u} = 29.05$$

bestimmt werden.

Aufgabe 10.1

Die Druckfestigkeit parallel zur Faser von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in der Tabelle gegeben.

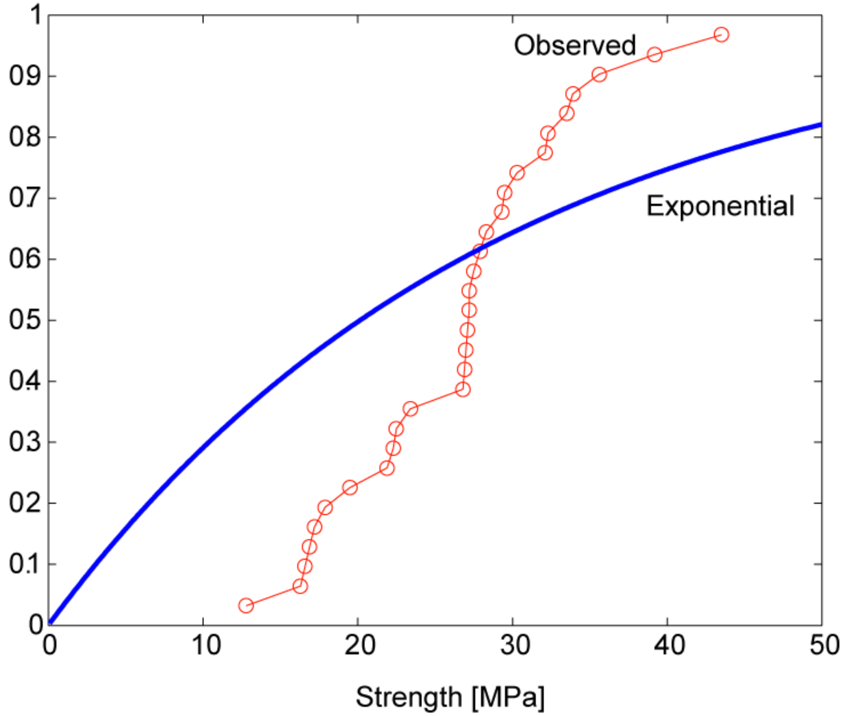
Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]
1	12.8	11	23.4	21	29.3
2	16.3	12	26.8	22	29.5
3	16.6	13	26.9	23	30.3
4	16.9	14	27	24	32.1
5	17.2	15	27.1	25	32.3
6	17.9	16	27.2	26	33.5
7	19.5	17	27.2	27	33.9
8	21.9	18	27.5	28	35.6
9	22.3	19	27.9	29	39.2
10	22.5	20	28.3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

b) Zeichne die kumulative Verteilungsfunktion der beiden Verteilungen und zeichnen Sie jeweils die kumulative Verteilung der Stichprobe ein.

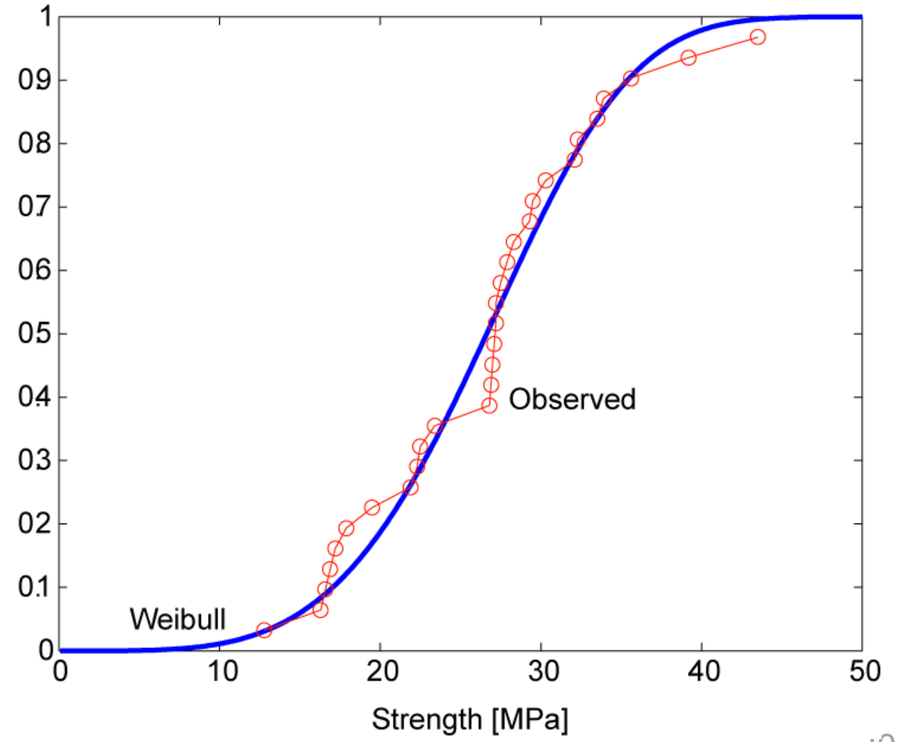
Aufgabe 10.1

b) Zeichne die kumulative Verteilungsfunktion der beiden Verteilungen und zeichne jeweils die kumulative Verteilung der Stichprobe ein.

Exponentialverteilung



Weibullverteilung



Aufgabe 10.1

Die Druckfestigkeit parallel zur Faser von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in der Tabelle gegeben.

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Intervall	Häufigkeit	Wahrscheinlichkeit P [Stichprobe in diesem Intervall]	Erwartete Häufigkeit
0-20			
20-25			
25-30			
30-∞			

Aufgabe 10.1

Exponentialverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutzedazu die Intervalle in der Tabelle.

Die Statistik für den Chi-Quadrat Test ist:

$$\varepsilon_{\mathrm{m}}^2 = \sum_{\mathrm{i}=1}^{\mathrm{k}} \frac{\left(N_i - n \cdot p_i\right)^2}{n \cdot p_i}$$

Und die Entscheidungsregel lautet dann:

$$P\left[\varepsilon^2 > c\right] = \alpha$$

$$P\left[\varepsilon^2 \leq c\right] = 1 - \alpha$$

ε^2 folgt der Chi-Quadrat Verteilung mit 2 FHG. Warum 2 Freiheitsgrade ?

Aufgabe 10.1

Exponentialverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

ε^2 folgt der Chi-Quadrat Verteilung mit 2 FHG.

Wir haben 4 Intervalle.
Das letzte Intervall ist abhängig von den 3 anderen – Reduktion um 1 FHG.
Bestimmung des Parameters λ aus den beobachteten Daten – Reduktion um 1 FHG.

4 – 1 – 1 = 2 FHG und $\alpha=0.1 \sim \chi = 4.6052$

f	$\chi^2_{F=0.01}$	$\chi^2_{F=0.05}$	$\chi^2_{F=0.10}$	$\chi^2_{F=0.25}$	$\chi^2_{F=0.50}$	$\chi^2_{F=0.75}$	$\chi^2_{F=0.90}$ $\alpha=10\%$
1	0.0002	0.0039	0.0158	0.1015	0.4549	1.3233	2.7055
2	0.0201	0.1026	0.2107	0.5754	1.3863	2.7726	4.6052
3	0.1148	0.3518	0.5844	1.2125	2.3660	4.1083	6.2514

Aufgabe 10.1

Exponentialverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Intervall	Häufigkeit N_i	Wahrscheinlichkeit P [Stichprobe in diesem Intervall]	Erwartete Häufigkeit	Normalisierte Quadrate der Differenzen
0-20	7	0.532		
20-25	4	0.081		
25-30	11	0.067		
30- ∞	8	0.32		
Summe	30	1		

beobachtete Werte N_1 im Intervall 1

$$\int_{20}^{25} \lambda \cdot e^{-\lambda x} dx$$

mit den zu $\lambda = 0.038$ geschätzten Parameter

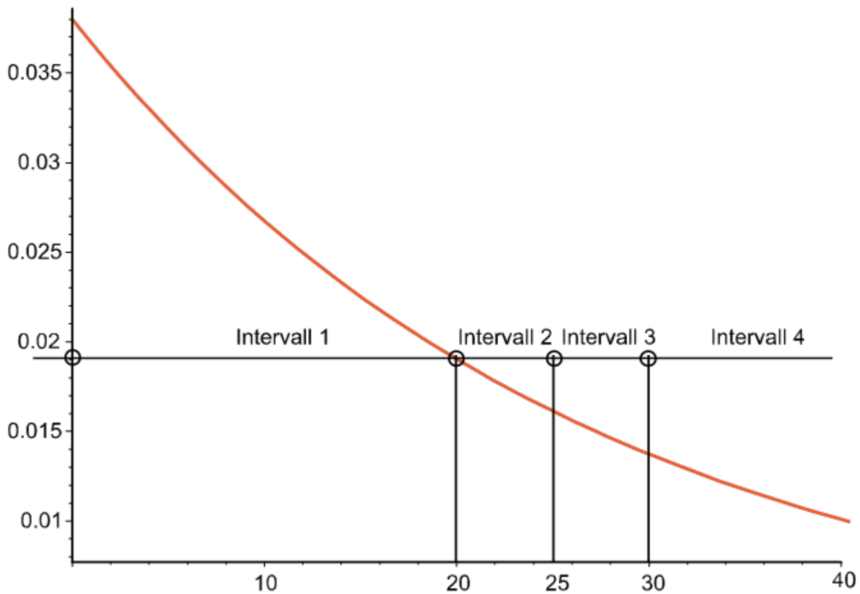
Aufgabe 10.1

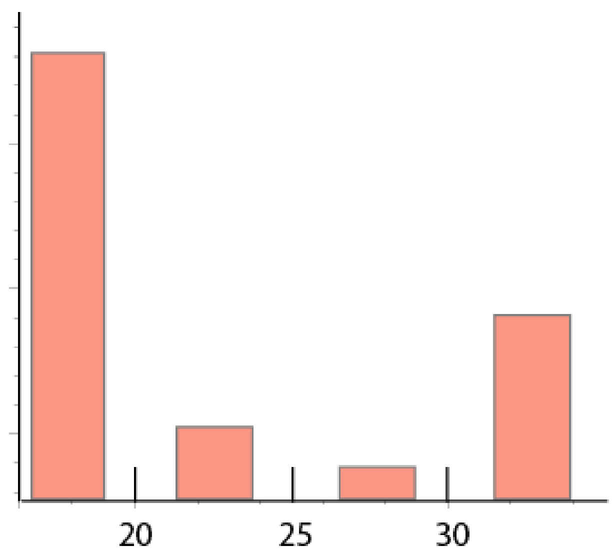
Exponentialverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Hier haben wir eine stetige Wahrscheinlichkeitsdichtefunktion gegeben und wollen den Chi-Quadrat Test anwenden.

Für den Test müssen wir deshalb die Wahrscheinlichkeitsdichtefunktion diskretisieren. – Hier in 4 Intervalle !





Aufgabe 10.1

Exponentialverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Intervall	Häufigkeit N_i	Wahrscheinlichkeit P [Stichprobe in diesem Intervall]	Erwartete Häufigkeit	Normalisierte Quadrate der Differenzen
0-20	7	0.532	15.96	5.03
20-25	4	0.081	2.43	1.01
25-30	11	0.067	2.01	40.21
30- ∞	8	0.32	9.60	0.27
Summe	30	1	30	46.52

beobachtete Werte N_1 im Intervall 1

$$\int_{20}^{25} \lambda \cdot e^{-\lambda x} dx$$

mit den zu $\lambda = 0.038$ geschätzten Parameter

$$n \cdot p_1 = 30 \cdot 0.532$$

$$\varepsilon_m^2 = \frac{(N_2 - n \cdot p_2)^2}{n \cdot p_2}$$

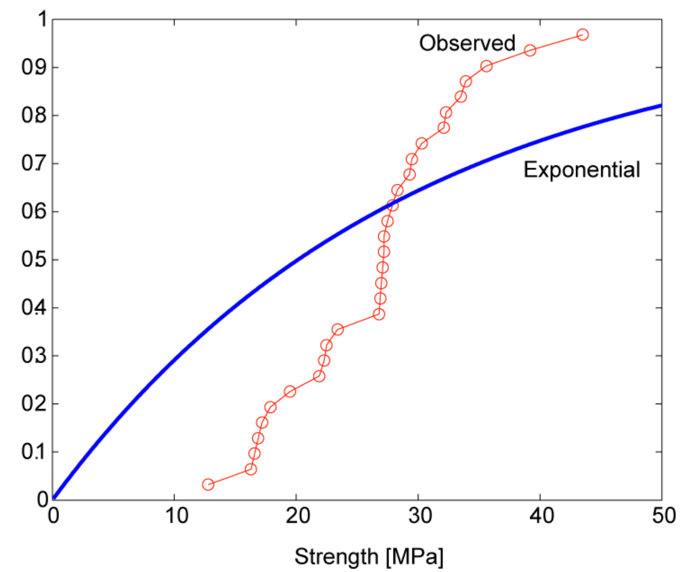
$$\varepsilon_m^2 = \sum_{i=1}^{k} \frac{(N_i - n \cdot p_i)^2}{n \cdot p_i}$$

Aufgabe 10.1

Exponentialverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

$$\varepsilon_m^2 = \sum_{i=1}^{k} \frac{\left(N_i - n \cdot p_i\right)^2}{n \cdot p_i} = 46.52$$



Mit der Entscheidungsregel :

$$P\left[\varepsilon^2 \leq c\right] = 1 - \alpha$$

Und dem aus der Tabelle ermittelten Wert für χ = 4.6052 folgt:

Da ε^2 =46.52 grösser als χ=4.61 ist, müssen wir die Hypothese, dass die Daten einer Exponentialverteilung folgen, auf einem Signifikanzniveau von 10% verwerfen.

Aufgabe 10.1

Weibullverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Die Statistik für den Chi-Quadrat Test ist:

$$\varepsilon_{\mathrm{m}}^2 = \sum_{\mathrm{i}=1}^{\mathrm{k}} \frac{\left(N_i - n \cdot p_i\right)^2}{n \cdot p_i}$$

Und die Entscheidungsregel lautet dann:

$$P\left[\varepsilon^2 > c\right] = \alpha$$

$$P\left[\varepsilon^2 \leq c\right] = 1 - \alpha$$

ε^2 folgt der Chi-Quadrat Verteilung mit 1 FHG.

Warum 1 Freiheitsgrad?

Aufgabe 10.1

Weibullverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

ε^2 folgt der Chi-Quadrat Verteilung mit 1 FHG.

Wir haben $m = 4$ Intervalle.
Das letzte Intervall ist abhängig von den 3 anderen – Reduktion um 1 FHG.
Bestimmung des Parameters k und u aus den beobachteten Daten – Reduktion um j=2 FHG.

$v = m - 1 - j$ = 4 – 1 – 2 = 1 FHG und a=0.1 ~ c = 2.7055

f	$\chi^2_{F=0.01}$	$\chi^2_{F=0.05}$	$\chi^2_{F=0.10}$	$\chi^2_{F=0.25}$	$\chi^2_{F=0.50}$	$\chi^2_{F=0.75}$	$\chi^2_{F=0.90}$ α=10%
1	0.0002	0.0039	0.0158	0.1015	0.4549	1.3233	2.7055
2	0.0201	0.1026	0.2107	0.5754	1.3863	2.7726	4.6052
3	0.1148	0.3518	0.5844	1.2125	2.3660	4.1083	6.2514

Aufgabe 10.1

Weibullverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Intervall	Häufigkeit	Wahrscheinlichkeit P [Stichprobe in diesem Intervall]	Erwartete Häufigkeit	Normalisierte Quadrate der Differenzen
0-20	7	0.19	5.7	0.296
20-25	4	0.22	6.6	1.024
25-30	11	0.27	8.1	1.038
30- ∞	8	0.32	9.6	0.267
Summe	30	1	30	2.63

beobachtete Werte N_1 im Intervall 1

$$\int_{20}^{25} \frac{k}{u} \cdot \left(\frac{x}{u}\right)^{k-1} \cdot e^{-\left(\frac{x}{u}\right)^k} dx$$

mit den zu $k = 4.21$ und $u = 29.05$ geschätzten Parameter

$n \cdot p_1 =$

$30 \cdot 0.19$

$$\varepsilon_m^2 = \frac{(N_2 - n \cdot p_2)^2}{n \cdot p_2}$$

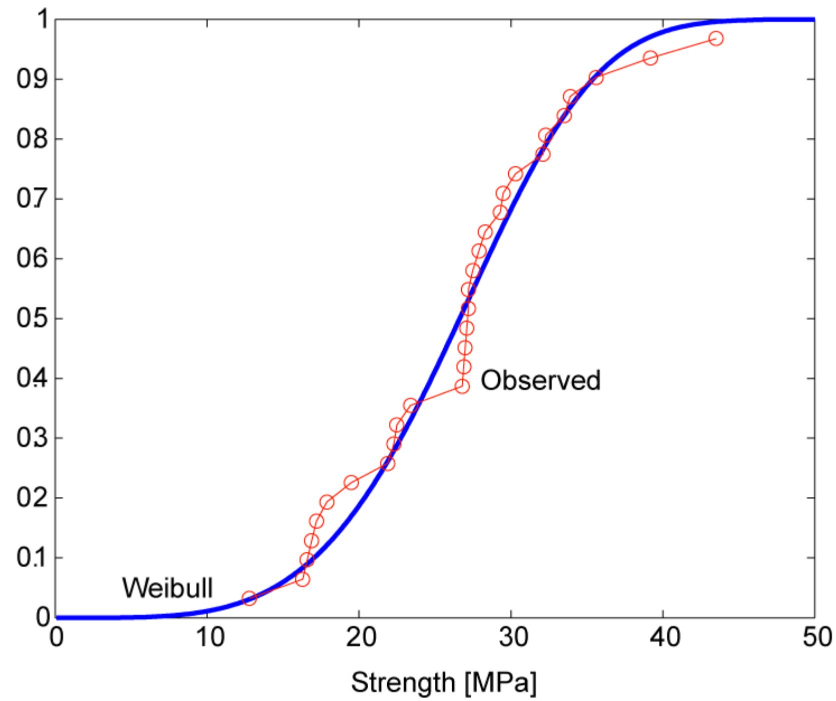
$$\varepsilon_m^2 = \sum_{i=1}^{k} \frac{(N_i - n \cdot p_i)^2}{n \cdot p_i}$$

Aufgabe 10.1

Weibullverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

$$\varepsilon_{\mathrm{m}}^2 = \sum_{\mathrm{i}=1}^{\mathrm{k}} \frac{\left(N_i - n \cdot p_i\right)^2}{n \cdot p_i} = 2.36$$



Mit der Entscheidungsregel :

$$P\left[\varepsilon^2 \leq c\right] = 1 - \alpha$$

Und den aus der Tabelle ermittelten Wert für χ = 2.71 folgt:

Da ε^2 =2.36 kleiner als χ=2.71 ist, können wir die Hypothese, dass die Daten einer Weibullverteilung folgen, auf einem Signifikanzniveau von 10% nicht verwerfen.

Aufgabe 10.1

Weibullverteilung

c) Teste die Güte der Anpassung für beide Verteilungen mit dem χ^2 Test auf einem Signifikanzniveau von 10%. Benutze dazu die Intervalle in der Tabelle.

Zu Beachten:

Der χ^2 Test hängt stark von der Wahl der Anzahl und der Grössen der Klassen ab.

Pro Klasse sollten mindestens 5 Beobachtungen enthalten, um ein plausibles Resultat zu erhalten.

Aufgabe 10.1

Die Druckfestigkeit parallel zur Faser von 30 Holzproben wurde bestimmt. Die Ergebnisse sind in der Tabelle gegeben.

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

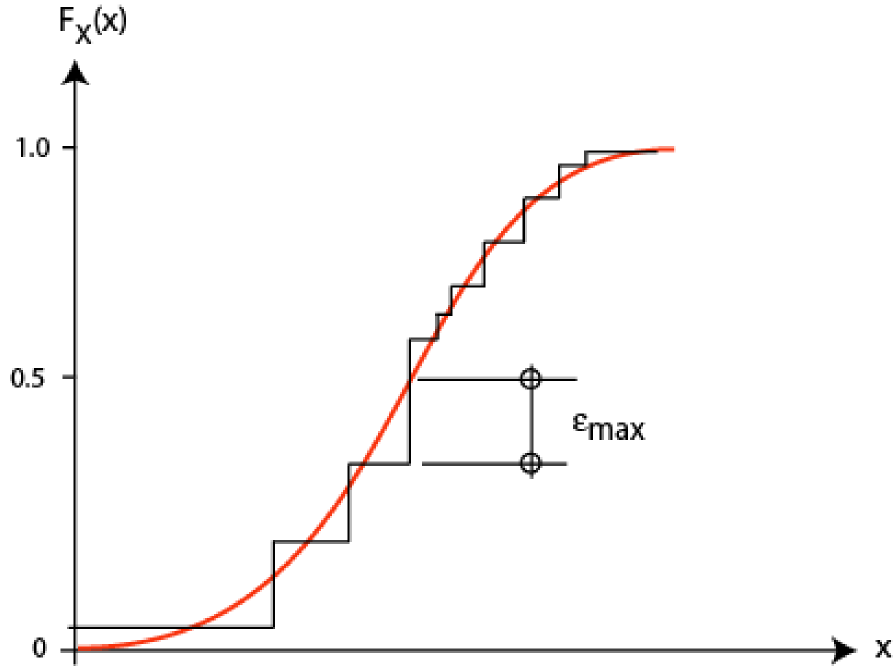
Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]	Nr.	Druck-festigkeit [MPa]
1	12.8	11	23.4	21	29.3
2	16.3	12	26.8	22	29.5
3	16.6	13	26.9	23	30.3
4	16.9	14	27	24	32.1
5	17.2	15	27.1	25	32.3
6	17.9	16	27.2	26	33.5
7	19.5	17	27.2	27	33.9
8	21.9	18	27.5	28	35.6
9	22.3	19	27.9	29	39.2
10	22.5	20	28.3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

Aufgabe 10.1

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Die Idee:

Bei diesem Test werden die Abweichungen der angenommenen kumulativen Verteilungsfunktion und der beobachteten kumulativen Verteilungsfunktion bestimmt. Die maximale Abweichung sollte dabei möglichst klein sein.



Aufgabe 10.1

d) Teste die Güte der Anpassung **beider** Verteilungen mit dem **Kolmogorov-Smirnov** Test auf einem Signifikanzniveau von 10%.

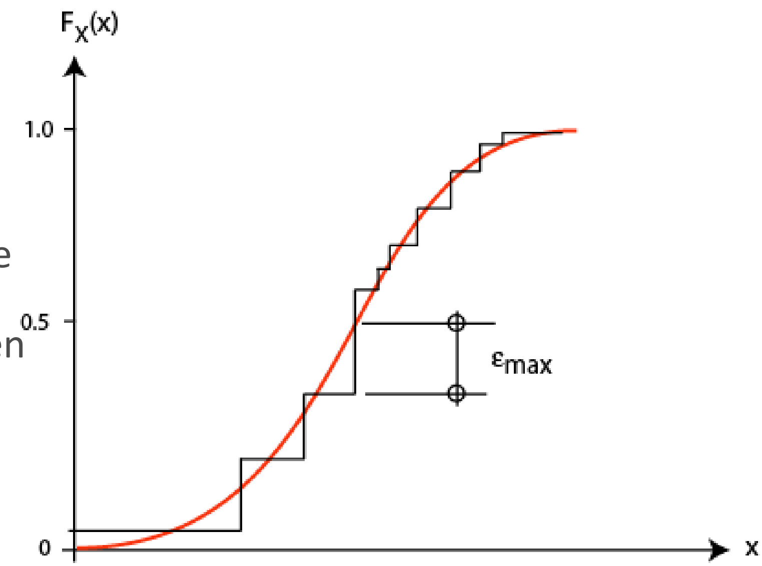
Dabei wird die beobachtete kumulative Verteilungsfunktion aus:

$$F_{o,X}(x) = \frac{i}{n}$$

i ← Rang in der geordneten Tabelle

n ← Gesamtzahl der Beobachtungen

berechnet.



Es ergibt sich die folgende Statistik:

$$\varepsilon_{\max} = \max_{i=1}^{n}\left[\left|F_{o,X}(x_i) - F_p(x_i)\right|\right] = \max_{i=1}^{n}\left[\left|\frac{i}{n} - F_p(x_i)\right|\right]$$

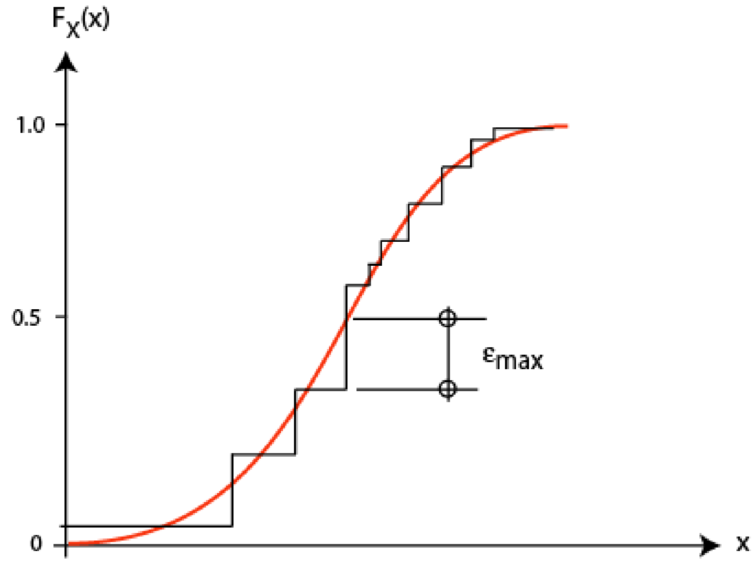
Aufgabe 10.1

d) Teste die Güte der Anpassung **beider** Verteilungen mit dem **Kolmogorov-Smirnov** Test auf einem Signifikanzniveau von 10%.

Und die Entscheidungsregel ergibt sich zu:

$$P[\varepsilon_{\max} > c] = \alpha$$

$$P[\varepsilon_{\max} \leq c] = 1 - \alpha$$



c kann für bestimmte Werte für a der Tabellen entnommen werden:

n	$\alpha = 0.01$	$\alpha = 0.02$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$
28	0.300	0.279	0.250	0.225	0.197
29	0.295	0.275	0.246	0.221	0.193
30	0.290	0.270	0.242	0.218	0.190
31	0.285	0.266	0.238	0.214	0.187

Aufgabe 10.1

Exponentialverteilung

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Die Stichprobenstatistik für den Kolmogorov-Smirnov-Test ist also

$$\varepsilon_{\max} = \max_{i=1}^{n}\left[\left|\frac{i}{n} - F_p(x_i)\right|\right]$$

Wobei $F_X(\ )$ die angenommene Verteilungsfunktion ist.

Schauen wir uns die Exponentialverteilung an:

$$F_X(x) = 1 - e^{-\lambda x}$$

Mit λ= 0.038 können wir die Werte berechnen

Aufgabe 10.1

Exponentialverteilung

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Nr.	[Mpa]	$F_X(x)$	i/n	$\mid i/n - F_X(x) \mid$
1	12.80	0.39	0.03	0.35
2	16.30	0.46	0.07	$\boxed{0.40}$
3	16.60	0.47	0.10	0.37
4	16.90	0.47	0.13	0.34
5	17.20	0.48	0.17	0.31
6	17.90	0.49	0.20	0.29
7	19.50	0.52	0.23	0.29
8	21.90	0.56	0.27	0.30
9	22.30	0.57	0.30	0.27
10	22.50	0.57	0.33	0.24
11	23.40	0.59	0.37	0.22
12	26.80	0.64	0.40	0.24
13	26.90	0.64	0.43	0.21
14	27.00	0.64	0.47	0.17
15	27.10	0.64	0.50	0.14
16	27.20	0.64	0.53	0.11
17	27.20	0.64	0.57	0.08
18	27.50	0.65	0.60	0.05
19	27.90	0.65	0.63	0.02
20	28.30	0.66	0.67	0.01
21	29.30	0.67	0.70	0.03
22	29.50	0.67	0.73	0.06
23	30.30	0.68	0.77	0.08
24	32.10	0.70	0.80	0.10
25	32.30	0.71	0.83	0.13
26	33.50	0.72	0.87	0.15
27	33.90	0.72	0.90	0.18
28	35.60	0.74	0.93	0.19
29	39.20	0.77	0.97	0.19
30	43.50	0.81	1.00	0.19

$$\max\left|\frac{i}{n} - FX(x)\right| = 0.40$$

$$F_X(x) = 1 - e^{-0.038x} = 1 - e^{-0.038\cdot 17.2} = 0.48$$

$$\frac{i}{n} = \frac{10}{30} = 0.333$$

$$\left|\frac{i}{n} - F_p(x_i)\right| =$$

$$|0.57 - 0.64| = 0.08$$

Aufgabe 10.1

Exponentialverteilung

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Mit der Entscheidungsregel können wir die Hypothese nun bewerten:

$$P\left[\varepsilon_{\max} \leq c\right] = 1 - \alpha$$

$\alpha = 0.1$ können wir den Wert *c* aus der Tabelle ablesen

n	$\alpha = 0.01$	$\alpha = 0.02$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$
28	0.300	0.279	0.250	0.225	0.197
29	0.295	0.275	0.246	0.221	0.193
30	0.290	0.270	0.242	0.218	0.190
31	0.285	0.266	0.238	0.214	0.187

c ergibt sich demnach zu 0.218

Da ε_{max} = 0.4 grösser ist als c=0.218, müssen wir die Hypothese, dass die Daten einer Exponentialverteilung folgen, auf einem Signifikanzniveau von 10% verwerfen.

Aufgabe 10.1

Weibullverteilung

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Da der Tabellenwert c unabhängig von der gewählten Verteilung und auch unabhängig vom Freiheitsgrad ist, gilt dieser auch für die Teststatistik der Weibullverteilung.

Die Stichprobenstatistik für den Kolmogorov-Smirnov-Test ist also

$$\varepsilon_{\max} = \max_{i=1}^{n}\left[\left|\frac{i}{n} - F_p(x_i)\right|\right]$$

Wobei $F_X(\ )$ die angenommene Verteilungsfunktion ist.

Schauen wir uns nun die Weibullverteilung an:

$$F_X(x) = 1 - e^{-\left(\frac{x}{u}\right)^k}$$

Mit u= 29.05 und k=4.21 können wir die Werte berechnen

Aufgabe 10.1

Weibullverteilung

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Nr.	[Mpa]	FX(x)	i/n	| i/n - FX(x) |
1	12.8	0.03	0.03	0.00
2	16.3	0.08	0.07	0.02
3	16.6	0.09	0.10	0.01
4	16.9	0.10	0.13	0.04
5	17.2	0.10	0.17	0.06
6	17.9	0.12	0.20	0.08
7	19.5	0.17	0.23	0.06
8	21.9	0.26	0.27	0.00
9	22.3	0.28	0.30	0.02
10	22.5	0.29	0.33	0.04
11	23.4	0.33	0.37	0.04
12	26.8	0.51	0.40	0.11
13	26.9	0.51	0.43	0.08
14	27	0.52	0.47	0.05
15	27.1	0.53	0.50	0.03
16	27.2	0.53	0.53	0.00
17	27.2	0.53	0.57	0.04
18	27.5	0.55	0.60	0.05
19	27.9	0.57	0.63	0.06
20	28.3	0.59	0.67	0.07
21	29.3	0.65	0.70	0.05
22	29.5	0.66	0.73	0.08
23	30.3	0.70	0.77	0.07
24	32.1	0.78	0.80	0.02
25	32.3	0.79	0.83	0.04
26	33.5	0.84	0.87	0.03
27	33.9	0.85	0.90	0.05
28	35.6	0.90	0.93	0.03
29	39.2	0.97	0.97	0.00
30	43.5	1.00	1.00	0.00

$$F_X(x) = 1 - e^{-\left(\frac{16.9}{29.05}\right)^{4.21}} = 0.10$$

$$\frac{i}{n} = \frac{10}{30} = 0.333$$

$$\max\left|\frac{i}{n} - FX(x)\right| = 0.11$$

$$\left|\frac{i}{n} - F_p(x_i)\right| =$$

$$|0.7 - 0.77| = 0.07$$

Aufgabe 10.1

Weibullverteilung

d) Teste die Güte der Anpassung beider Verteilungen mit dem Kolmogorov-Smirnov Test auf einem Signifikanzniveau von 10%.

Mit der Entscheidungsregel können wir die Hypothese nun beurteilen:

$$P\left[\varepsilon_{\max} \leq c\right] = 1 - \alpha$$

Da ε_{max} = 0.11 kleiner ist als c=0.218, können wir die Hypothese, dass die Daten einer Weibullverteilung folgen, auf einem Signifikanzniveau von 10% annehmen.

Aufgabe 10.2

Anhand eines Teiles des in der ersten Vorlesung erhobenen Datensatzes, welcher die Körpergrösse aller Frauen beinhaltet, soll folgendes durchgeführt werden:

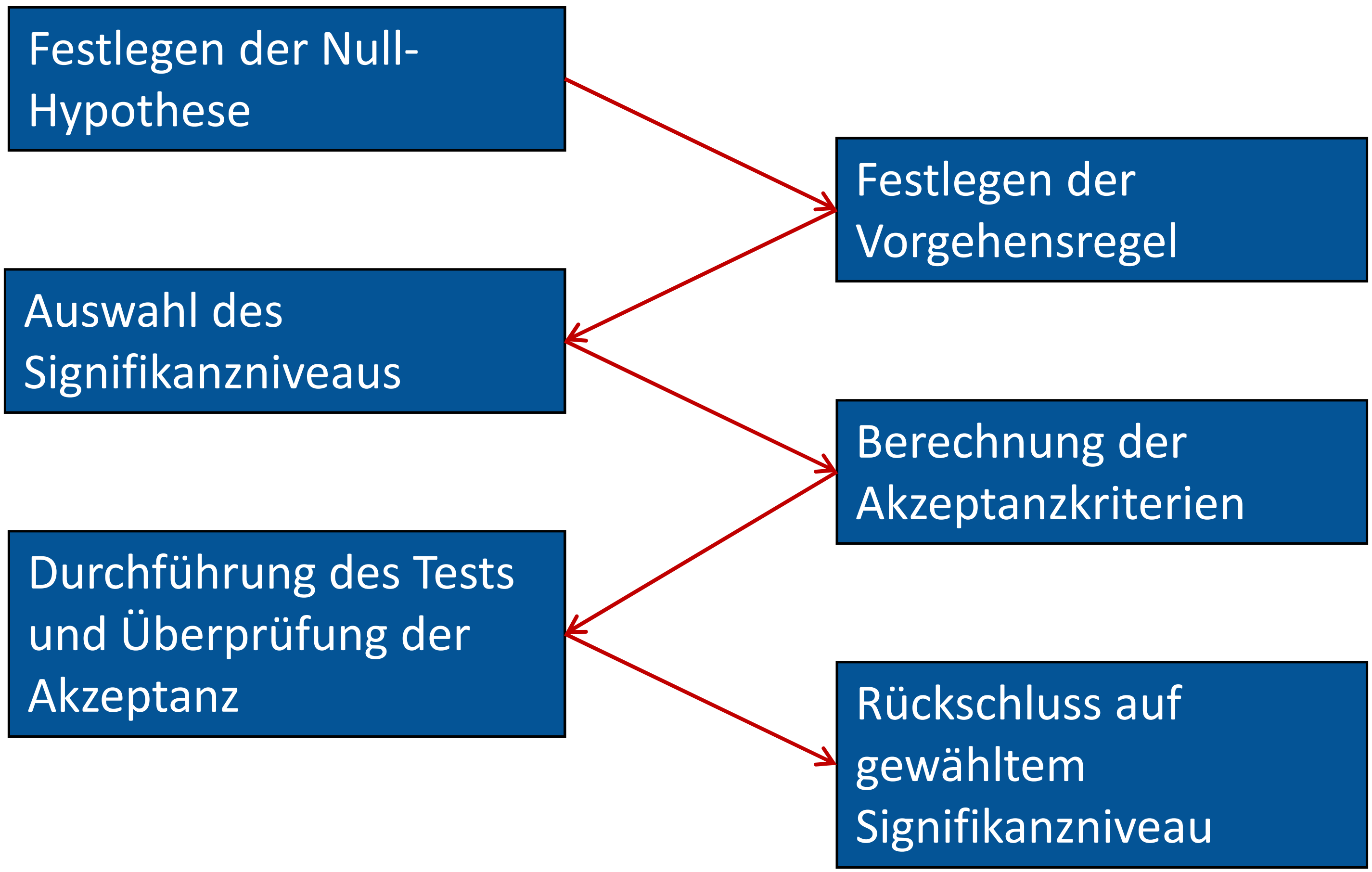
a) Überprüfe mittels eines T-Tests ob, der Mittelwert einem gegebenen Literaturwert von μ = 168 cm entspricht, auf einem Signifikanzniveau von 20%.

b) Passe die Normalverteilung den Beobachtungen an. Bestimme dazu die Parameter μ und σ mit der Maximum Likelihood Methode.

a) Teste die Güte der Anpassung für die Normalverteilung mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

Testen von Hypothesen

- Generelles Vorgehen beim Hypothesentest



Aufgabe 10.2

a) Überprüfe mittels eines T-Tests, ob der Mittelwert einem gegebenen Literaturwert von μ = 168 cm entspricht, auf einem Signifikanzniveau von 20%.

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		

$$x_M = 168.92 cm$$
$$S_M = 6.41 cm$$

Aufgabe 10.2

a) Überprüfe mittels eines T-Tests, ob der Mittelwert der Strichprobe μ_M einem gegebenen Literaturwert von μ_L = 168 cm entspricht, auf einem Signifikanzniveau von 20%.

1. Festlegen der Nullhypothese

$$H_0 : \mu_M = 168cm$$
$$H_1 : \mu_M \neq 168cm$$

2. Festlegen der Entscheidungsregel

Der Stichprobenmittelwert muss in einem zu bestimmenden Intervall rund um den Literaturmittelwert μ_L liegen.

$$P[-t_{\alpha/2} \leq T = \frac{\bar{X}_M - \mu_L}{S_M / \sqrt{n}} \leq t_{\alpha/2}] = 1 - \alpha$$

Für die Stichprobe wird dies berechnet als:

$$\mu_L - t_{\alpha/2} \frac{s_M}{\sqrt{n}} \leq \bar{x}_M \leq \mu_L + t_{\alpha/2} \frac{s_M}{\sqrt{n}}$$

Aufgabe 10.2

$$\bar{x}_M = 168.92cm$$
$$S_M = 6.41cm$$

a) Überprüfe mittels eines T-Tests, ob der Mittelwert der Strichprobe μ_M einem gegebenen Literaturwert von μ_L = 168 cm entspricht, auf einem Signifikanzniveau von 20%.

3. Auswahl des Signifikanzniveaus: α = 20 %

4. Berechnung der Akzeptanzkriterien

Intervall für den Stichprobenmittelwert berechnen:

$$\mu_L - t_{\alpha/2}\frac{s_M}{\sqrt{n}} \leq \bar{x}_M \leq \mu_L + t_{\alpha/2}\frac{s_M}{\sqrt{n}}$$

Untere Intervallgrenze: $\mu_L - t_{\alpha/2}\frac{s_M}{\sqrt{n}} = 168 - 1.303 \cdot \frac{6.41}{\sqrt{38}} = 166.65$

Obere Intervallgrenze: $\mu_L + t_{\alpha/2}\frac{s_M}{\sqrt{n}} = 168 + 1.303 \cdot \frac{6.41}{\sqrt{38}} = 169.35$

Aufgabe 10.2

a) Überprüfe mittels eines T-Tests, ob der Mittelwert der Strichprobe μ_M einem gegebenen Literaturwert von μ_L = 168 cm entspricht, auf einem Signifikanzniveau von 20%.

5. Durchführung des Tests

Der Stichprobenmittelwert wurde berechnet zu: $\bar{x}_M = 168.92cm$

Das Intervall der t-Statistik wurde berechnet als [166.65cm; 169.35cm].

Der Stichprobenmittelwert liegt im Intervall.

6. Rückschluss auf gewähltem Signifikanzniveau

Die Nullhypothese, dass es sich um die gleichen Populationen handelt, wird auf einem Signifikanzniveau von 20% angenommen.

Aufgabe 10.2

b) Passe die Normalverteilung den Beobachtungen an. Bestimme dazu die Parameter μ und σ mit der Maximum Likelihood Methode.

Normalverteilung:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Likelihood:

$$L(\boldsymbol{\theta}|\hat{\mathbf{x}}) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\hat{x}_i-\mu}{\sigma}\right)^2\right)$$

Gegebener Parameter (Literaturwert):

$$\theta_1 = \mu_L = 168$$

Zu Bestimmender Parameter:

$$\theta_2 = \hat{\sigma}_X = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - \theta_1)^2}{n}} = 10.934$$

Aufgabe 10.2

$$\mu = 168cm$$

$$\hat{\sigma} = 10.934cm$$

c) Teste die Güte der Anpassung für die Normalverteilung mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

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 10.1

Exponentialverteilung

c) Teste die Güte der Anpassung für die Normalverteilung mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

ε^2 folgt der Chi-Quadrat Verteilung mit 3 FHG.

Wir haben 5 Intervalle
Das letzte Intervall ist abhängig von den 4 anderen – Reduktion um 1 FHG.
Bestimmung des Parameters σ aus den beobachteten Daten – Reduktion um 1 FHG.

5– 1 – 1 = 3 FHG und α=0.1 ~ χ = 6.2514

f	$\chi^2_{F=0.01}$	$\chi^2_{F=0.05}$	$\chi^2_{F=0.10}$	$\chi^2_{F=0.25}$	$\chi^2_{F=0.50}$	$\chi^2_{F=0.75}$	$\chi^2_{F=0.90}$ α=10%
1	0.0002	0.0039	0.0158	0.1015	0.4549	1.3233	2.7055
2	0.0201	0.1026	0.2107	0.5754	1.3863	2.7726	4.6052
3	0.1148	0.3518	0.5844	1.2125	2.3660	4.1083	6.2514

Aufgabe 10.2

$$\mu = 168cm$$

$$\hat{\sigma} = 10.934cm$$

c) Teste die Güte der Anpassung für die Normalverteilung mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

Intervall	Häufigkeit	Wahrscheinlichkeit P [Stichprobe in diesem Intervall]	Erwartete Häufigkeit	Normalisierte Quadrate der Differenzen
0-160	5	0.2322		
161-165	9	0.1597		
166-170	8	0.1807		
171-175	10	0.1664		
176- ∞	6	0.2610		
Summe	38	1		

$$\frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_{161}^{165} e^{\left(-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2\right)} dx$$

Aufgabe 10.2

$$\mu = 168cm$$

$$\hat{\sigma} = 10.934cm$$

c) Teste die Güte der Anpassung für die Normalverteilung mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

Intervall	**Häufigkeit**	**Wahrscheinlichkeit P [Stichprobe in diesem Intervall]**	**Erwartete Häufigkeit**	**Normalisierte Quadrate der Differenzen**
0-160	5	0.2322	8.824	1.657
161-165	9	0.1597	6.069	1.418
166-170	8	0.1807	6.867	0.1869
171-175	10	0.1664	6.323	2.1383
176- ∞	6	0.2610	9.918	1.548
Summe	38	1	38	$\boxed{6.948}$

$$\varepsilon_m^2 = \frac{(N_2 - n \cdot p_2)^2}{n \cdot p_2}$$

$$\frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_{161}^{165} e^{\left(-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2\right)} dx$$

$$n \cdot p_1 = 38 \cdot 0.2322$$

$$\varepsilon_m^2 = \sum_{i=1}^{k} \frac{(N_i - n \cdot p_i)^2}{n \cdot p_i}$$

Aufgabe 10.2

$$\mu = 168cm$$

$$\hat{\sigma} = 10.934cm$$

c) Teste die Güte der Anpassung für die Normalverteilung mit dem χ^2 Test auf einem Signifikanzniveau von 10%.

Mit dem aus der Tabelle ermittelten Wert für χ = 6.2514 folgt:

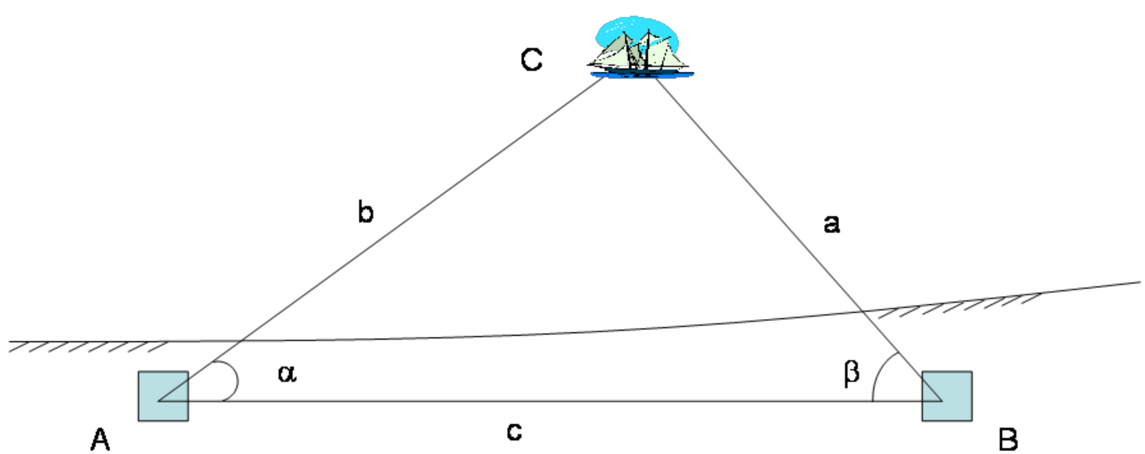
Da ε^2 =6.948 grösser als χ= 6.2514 ist, können wir die Hypothese, dass die Daten dieser Normalverteilung folgen, auf einem Signifikanzniveau von 10% nicht annehmen.

Aufgabe 10.3 Gruppenaufgabe

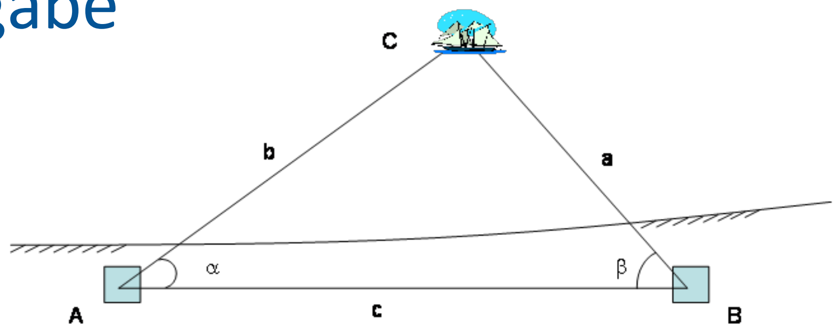
Die Position eines Schiffs kann von zwei Punkten A und B aus, welche sich auf dem Festland befinden, eindeutig bestimmt werden. Die Winkel α und β werden von der Basislinie AB aus gemessen.

Bestimme den Fehler in b, wenn die folgenden Daten bekannt sind:

$$c = 6\ km \pm 0.005\ km$$
$$\alpha = 0.813\ rad \pm 0.011\ rad$$
$$\beta = 1.225\ rad \pm 0.011\ rad$$



Aufgabe 10.3 Gruppenaufgabe



b ist durch α, β und c bestimmt:

$$b = \frac{\sin\beta}{\sin(\alpha+\beta)}c$$

Die Frage lautet also: Wie kann die Unsicherheit in b aus den Unsicherheiten in α, β und c berechnet werden?

$$c = 6\ km \pm 0.005\ km$$
$$\alpha = 0.813\ rad \pm 0.011\ rad \xrightarrow{b = \frac{\sin\beta}{\sin(\alpha+\beta)}c} b = ???\, km \pm ???\, km$$
$$\beta = 1.225\ rad \pm 0.011\ rad$$

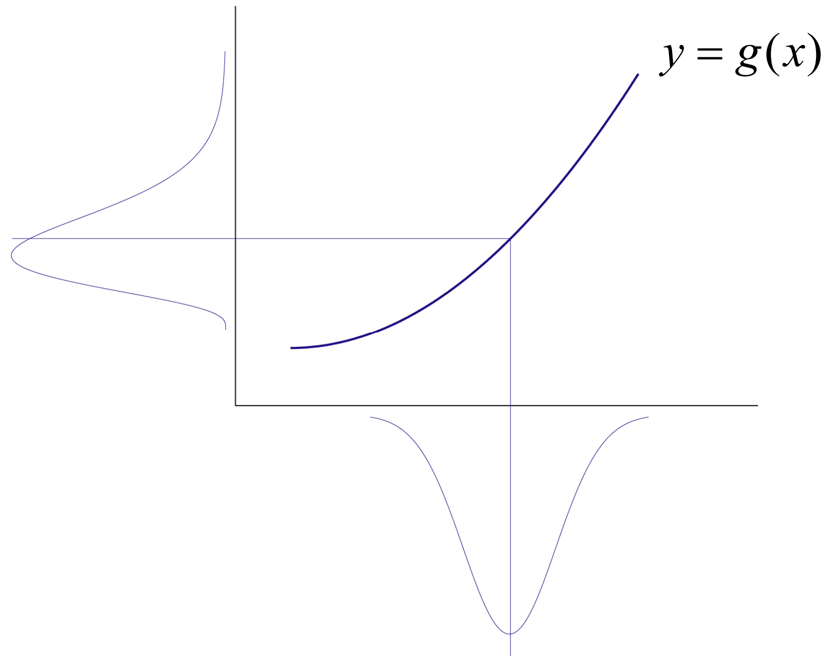
Oder allgemein:

$$\text{Unsicherheit X} \xrightarrow{Y = g(X)} \text{Unsicherheit Y}$$

Aufgabe 10.3 Gruppenaufgabe

Das Fehlerfortpflanzungsgesetz:

Wie überträgt sich die Unsicherheit in X auf die Unsicherheit in $Y = g(X)$?

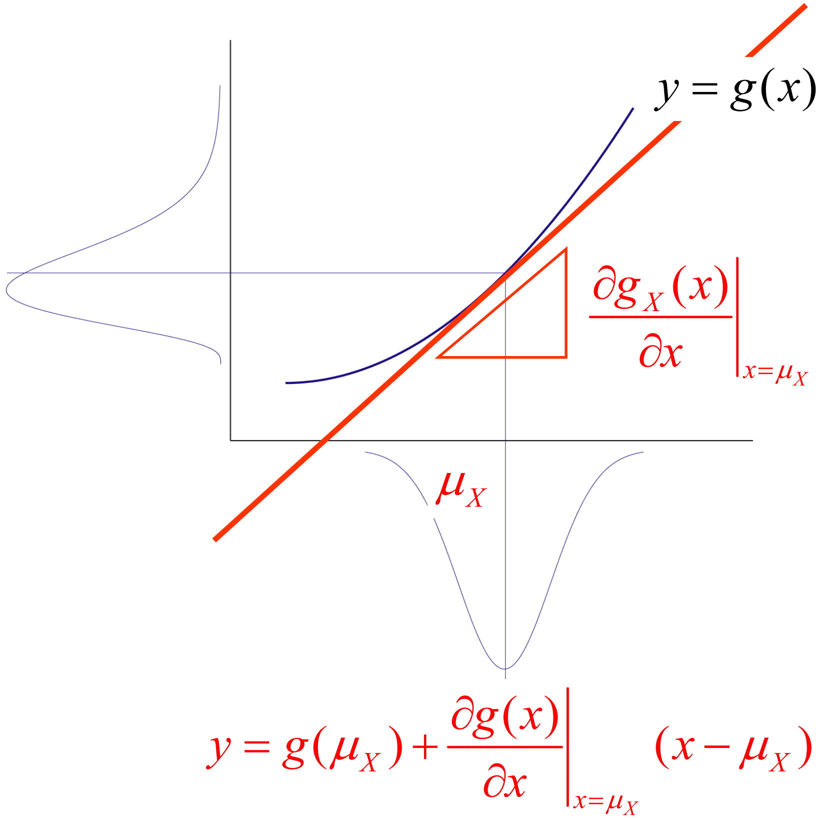


Bei einer Transformation wird die Dichtefunktion verändert…

Wie verhält sich dabei der Mittelwert?
Was passiert mit der Standardabweichung?

Aufgabe 10.3(12.1)

Das Fehlerfortpflanzungsgesetz:



Zur Erinnerung:

$E[aX+b]=aE[X]+b$
$Var[aX+b]=a^2 Var[X]$

Das bedeutet:

Wenn Y linear von X abhängt,
kann der Mittelwert von Y aus
demjenigen von X berechnet werden.

Der Zusammenhang zwischen X und Y kann wie folgt angenähert werden:

$$Y = g(\mu_X) + \left.\frac{\partial g(X)}{\partial x}\right|_{X=\mu_X} (X - \mu_X)$$