

1. Teilprüfung Statistik und Wahrscheinlichkeitsrechnung

FS 2010

Prof. Dr. Michael Havbro Faber

ETH Zürich

1.4. 2010
08:00 – 09:30

Vorname:

Name:

Stud. Nr.:

Studienrichtung:

1. Teilprüfung: Statistik und Wahrscheinlichkeitsrechnung Bau-, Umwelt- und Geomatikingenieurwissenschaften

Datum und Dauer:

Donnerstag, 1. April 2009

Beginn: 8:00 Uhr

Zeitdauer: 90 Minuten

Hilfsmittel:

- Alle Unterlagen (Skripte, Bücher, andere Ausdrücke, etc.) sind erlaubt.
- Taschenrechner (ohne Kommunikationsmittel) sind erlaubt.
- Kommunikationsmittel (z.B. Telefon) sind nicht erlaubt.

Hinweise:

- Bitte kontrollieren Sie zuerst, ob Sie das Material vollständig erhalten haben:
 - Aufgabenstellung inkl. genereller Information (18 Seiten).
 - Papierbogen kariert und gestempelt (1 mal).
- Bitte legen Sie Ihre Legi vor sich auf den Tisch.
- Alle Lösungsblätter müssen mit Namen und Vornamen versehen werden.
- Nur die zur Verfügung gestellten Blätter dürfen verwendet werden.
- Legen Sie am Ende der Prüfung alle Aufgaben- und Lösungsblätter in das Couvert zurück und lassen Sie dieses am Platz liegen.
- Wenn Sie vor 9:00 Uhr fertig sind, dann benachrichtigen Sie einen Assistierenden; er/sie wird dann Ihre Prüfung einsammeln. Sie dürfen bis 9:00 Uhr den Saal verlassen; danach warten Sie bitte, bis die Prüfung zu Ende ist (9:30 Uhr).

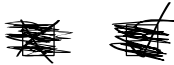
Teil 1: Multiple Choice (maximal 54 Punkte)

In den folgenden Multiple Choice Fragen können für die gleiche Frage (mindestens) eine oder mehrere Antworten zutreffend sein.

Bitte markieren Sie alle richtigen Antworten mit einem Häkchen oder Kreuz:



Wenn Sie ein bereits markiertes Kästchen rückgängig machen wollen, dann tun Sie das bitte deutlich:



Auf jede Aufgabe bzw. Teilaufgabe werden maximal 2 Punkte vergeben.

1.1 In welcher Phase / in welchen Phasen des Lebenszyklus eines Gebäudes spielen Unsicherheiten eine Rolle?

- Konzept, Machbarkeitsstudie
- Planung und Bemessung
- Nutzungsphase
- Rückbau der Anlage

1.2 Zwei alternative Konstruktionsweisen einer Brücke sind mit unterschiedlichen Wahrscheinlichkeiten für einen Einsturz verbunden:

Alternative 1: $P(\text{Einsturz}) = 10^{-6} / \text{Jahr}$

Alternative 2: $P(\text{Einsturz}) = 10^{-7} / \text{Jahr}$

Welche der folgenden Aussagen ist/sind korrekt, wenn die Kosten im Falle eines Einsturzes für beide Alternativen gleich sind?

- Das mit den beiden Alternativen verbundene Risiko ist gleich.
- Das mit Alternative 1 verbundene Risiko beträgt $10^{-6} / \text{Jahr}$.
- Das mit Alternative 1 verbundene Risiko ist 10 mal höher als das von Alternative 2.
- Das mittlere Einsturzrisiko der Brücke ist $5 \cdot 10^{-7}$.

1.3 Welche der folgenden Aussagen ist/sind korrekt?

Vertreter der frequentistischen Wahrscheinlichkeitsdefinition bestimmen Wahrscheinlichkeiten mit Hilfe von Experimenten.

Für „Bayesianer“ sind Wahrscheinlichkeiten objektiv eindeutig bestimmbar.

Der Satz von Bayes ist nur bei Annahme der Bayes'schen Interpretation der Wahrscheinlichkeit richtig.

Die Bayes'sche Interpretation der Wahrscheinlichkeit umfasst sowohl die frequentistische als auch die klassische Interpretation.

1.4 Für welche der folgenden Wahrscheinlichkeiten ergibt die Berechnung auf Basis von frequentistischer Information sinnvolle Ergebnisse?

Die Wahrscheinlichkeit, dass eine neue Leuchtstoffröhre beim ersten Einschalten defekt ist.

Die Wahrscheinlichkeit, dass die Festigkeit eines S235 Baustahls 235 N/mm^2 unterschreitet.

Die jährliche Wahrscheinlichkeit, dass die Züricher Hardbrücke einstürzt.

Die jährliche Wahrscheinlichkeit, dass die Züricher Hardbrücke durch eine Erdbebenbelastung einstürzt.

1.5 Welche der folgenden Aussagen gehört / gehören zu den drei grundlegenden Axiomen der Wahrscheinlichkeitstheorie?

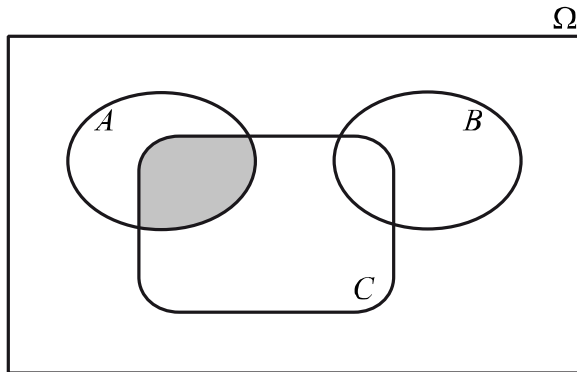
Die Wahrscheinlichkeit des sicheren Ereignisses ist Eins.

Für zwei unvereinbare Ereignisse E_1 und E_2 lässt sich die Wahrscheinlichkeit der Schnittmenge der beiden Ereignisse $P(E_1 \cap E_2)$ als Produkt der marginalen Wahrscheinlichkeiten $P(E_1)$ und $P(E_2)$ berechnen.

Die Schnittmenge von zwei unabhängigen Ereignissen ist die leere Menge.

Wahrscheinlichkeiten sind reelle Zahlen zwischen Null und Eins.

- 1.6 In der folgenden Abbildung sind drei Ereignisse A , B und C im Ereignisraum Ω dargestellt:



Welche der folgenden Aussagen ist/sind korrekt?

- Die Ereignisse A und B sind voneinander unabhängig.
- Die Vereinigung der Ereignisse A und C ist die leere Menge.
- Die graue Fläche stellt die Schnittmenge von A und C dar.
- In der Mengenschreibweise heisst die graue Fläche $A \cup C$.

- 1.7 Die Ereignisse A und B sind unvereinbar (schliessen sich gegenseitig aus). Welche der folgenden Gleichungen ist/sind dann korrekt?

- $P(A \cap B) = P(A) \cdot P(B)$
- $P(A \cup B) = P(A) + P(B)$
- $P(\bar{A} \cap B) = P(\bar{A} | B) \cdot P(B)$
- $P(\bar{A} \cup \bar{B}) = P(\bar{A}) + P(\bar{B}) - P(\bar{A} \cap \bar{B})$

- 1.8 Der Brand einer Berghütte kann entweder durch einen Blitzschlag ausgelöst werden (Ereignis B) oder durch eine andere Zündquelle (Ereignis A). Die beiden Ereignisse sind unabhängig voneinander und die folgenden jährlichen Eintrittswahrscheinlichkeiten sind bekannt:

$$P(B) = 0.01 \quad \text{und} \quad P(A) = 0.02$$

Welche der folgenden Aussagen ist/sind korrekt?

Die jährliche Brandeintrittswahrscheinlichkeit beträgt 0.0002.

Die Wahrscheinlichkeit, dass es in der Hütte im Jahr 2010 genau einmal brennt (egal welche Ursache), beträgt $0.01 + 0.02 = 0.03$.

Die Wahrscheinlichkeit, dass die Hütte im selben Jahr sowohl durch einen Blitzschlag, als auch (ein zweites Mal) durch eine andere Ursache in Brand gerät, beträgt 0.0002.

Die Wahrscheinlichkeit, dass es im Jahr 2010 in der Berghütte gar nicht brennt, beträgt $0.99 \cdot 0.98 = 0.9702$.

- 1.9 Für das nächste Richtfest kauft der Zimmermanns-Lehrling 12 Käsebrötchen und 8 Wurstbrötchen. Zum Spass bestreicht er 5 der Käse- und 5 der Wurstbrötchen unter dem Belag mit Chilipaste. Welche der folgenden Aussagen ist / sind korrekt?

Die Wahrscheinlichkeit, dass ein beliebiges Wurstbrötchen mit Chilipaste bestrichen ist, beträgt $5/8$.

Die Wahrscheinlichkeit, dass ein beliebiges Käsebrötchen mit Chilipaste bestrichen ist, berechnet sich zu $5/(12+8)$.

Die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Brötchen (Wurst oder Käse) mit Chilipaste bestrichen ist, berechnet sich zu $\frac{12}{20} \cdot \frac{5}{12} + \frac{8}{20} \cdot \frac{5}{8}$.

Die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Brötchen (Wurst oder Käse) mit Chilipaste bestrichen ist, berechnet sich zu $(5+5)/(12+8)$.

- 1.10 Für zwei sich gegenseitig ausschliessende wahre Zustände E_1 und E_2 und eine Beobachtung A lässt sich der Satz von Bayes wie folgt schreiben:

$$P(E_1|A) = \frac{P(A|E_1) \cdot P(E_1)}{P(A|E_1) \cdot P(E_1) + P(A|E_2) \cdot P(E_2)}$$

Welche der folgenden Aussagen ist/sind korrekt?

Der Nenner $P(A|E_1) \cdot P(E_1) + P(A|E_2) \cdot P(E_2)$ lässt sich aus dem Satz der totalen Wahrscheinlichkeit herleiten.

Die *a priori* Wahrscheinlichkeit $P(E_1)$ ist bereits vor dem Erhalt der Information über das Ereignis A bekannt.

Für die Likelihoods gilt $P(A|E_1) + P(\bar{A}|E_1) = 1$.

Die *a posteriori* Wahrscheinlichkeit $P(E_1|A)$ bezeichnet die Wahrscheinlichkeit einer Beobachtung A gegeben den wahren Zustand E_1 .

- 1.11 Welche der folgenden Aussagen ist/sind korrekt?

Der Variationskoeffizient einer Stichprobe ist dimensionslos.

Um die Streuung von zwei Datensätzen mit unterschiedlichen Dimensionen (Einheiten) miteinander zu vergleichen, eignet sich vor allem der Stichproben-Median.

Der Mittelwert einer Stichprobe entspricht dem am häufigsten vorkommenden Wert.

Der Schiefekoeffizient ist ein geeignetes Mass für die Streuung eines Datensatzes.

- 1.12 Welche der folgenden Aussagen trifft/treffen für eine symmetrische Häufigkeitsverteilung zu, die nur einen Modus hat?

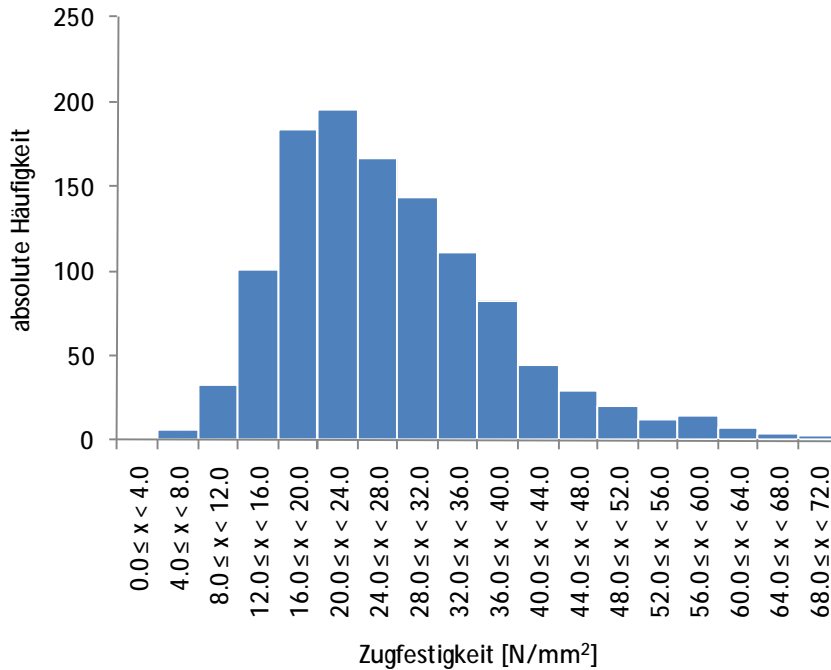
Die Standardabweichung entspricht der Wurzel aus der Varianz.

Der Stichproben-Mittelwert entspricht dem Modus.

Der Median entspricht dem Modus.

Die Kurtosis ist gleich Null.

1.13 Gegeben ist das folgende Histogramm mit den relativen Häufigkeiten der Zugfestigkeit von Schnittholz einer Stichprobe mit der Grösse $n=1162$.



Welche der folgenden Aussagen ist/sind richtig?

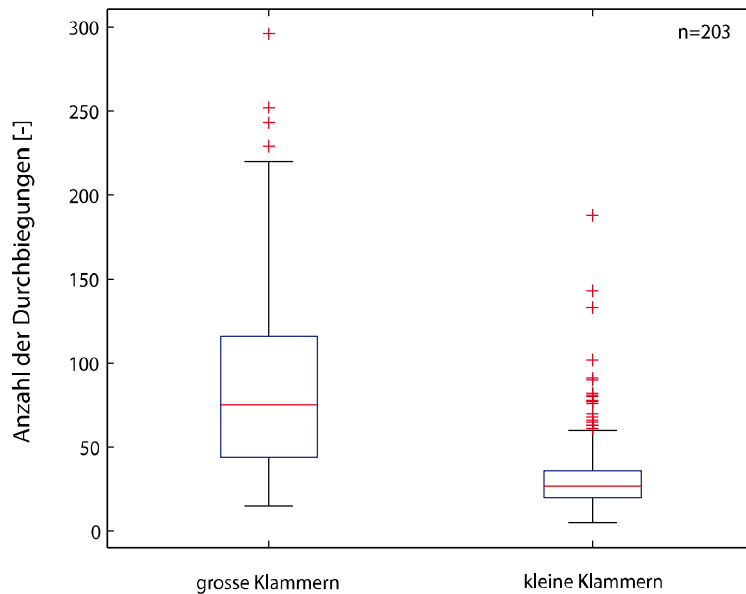
Der Modus hat einen Wert von 24 N/mm².

Der Modus entspricht dem Intervall [20-24) N/mm².

Der Mittelwert hat einen niedrigeren Wert als der Modus.

Das Histogramm ist rechtschief.

1.14 Die folgende Abbildung zeigt zwei Tukey-Box-Plots für die Anzahl Durchbiegung bis zum Versagen von grossen und kleinen Büroklammern.



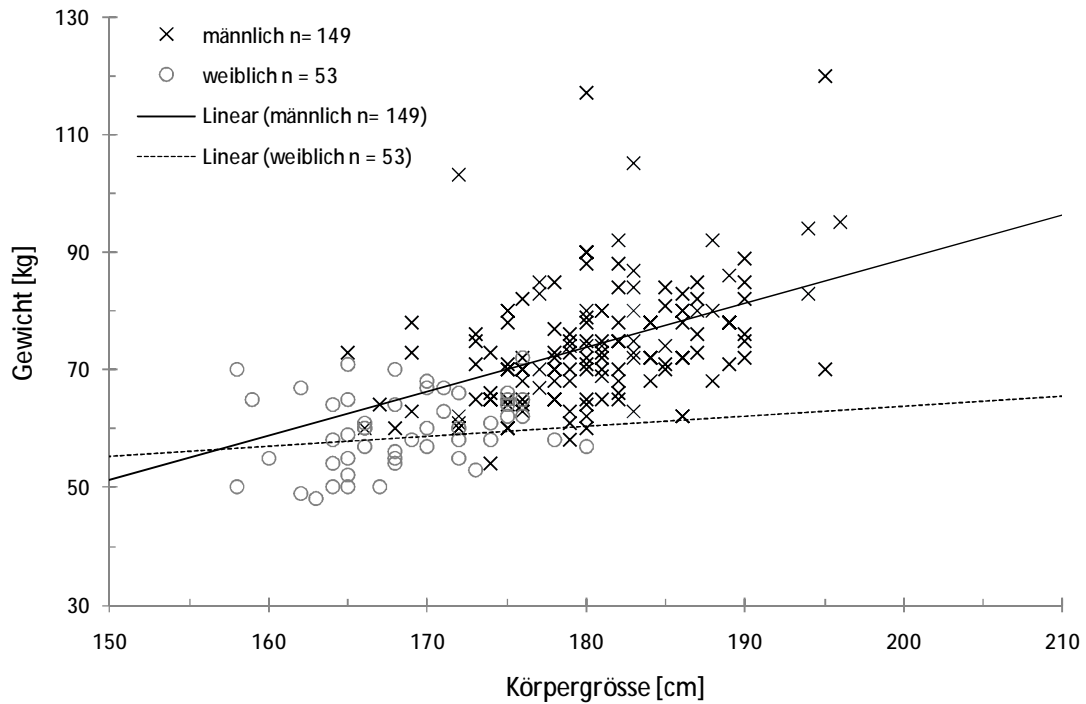
Aus der Abbildung ist erkennbar, dass

- alle aus dem Tukey Box Plot für die grossen Klammern ablesbaren Werte höher sind als bei den kleinen Klammern.
- die Häufigkeitsverteilung bei den kleinen Klammern linksschief ist.
- die Werte der grossen Klammern stärker um ihren Median streuen als die der kleinen Klammern.
- der Tukey-Box-Plot ein sehr gutes Instrument ist, um die Korrelation der beiden Datenreihen aufzuzeigen.

1.15 Ein Wert \hat{x}_i wird bei einem Tukey-Box-Plot dann als Ausreisser deklariert, wenn

- $\hat{x}_i > Q_{0.75} - 1.5 \cdot r$
- $\hat{x}_i > Q_{0.75} + 1.5 \cdot r$
- $\hat{x}_i > Q_{0.25} + 1.5 \cdot r$
- $\hat{x}_i < Q_{0.25} - 1.5 \cdot r$

1.16 Die folgende Abbildung zeigt ein zweidimensionales Streudiagramm, in dem das Körpergewicht der Studierenden über deren Körpergrösse getrennt nach Geschlecht aufgetragen ist.



Aus der Abbildung ist erkennbar, dass

- der Korrelationskoeffizient zwischen der Körpergrösse und dem Gewicht für Männer positiv ist.
- mit zunehmender Körpergrösse der Frauen auch das Gewicht der Männer zunimmt.
- der Zusammenhang zwischen Gewicht und Körpergrösse bei den Frauen ungefähr einem Korrelationskoeffizienten von $r = 0.9$ entspricht.
- der Stichprobenraum für das Gewicht der Frauen einen Bereich von ca. 49kg bis ca. 71kg umspannt.

1.17 Ein Korrelationskoeffizient

- von $r=1$ sagt aus, dass die Datenpaare zweier Datensätze perfekt linear und positiv miteinander korrelieren.
- von $r=0$ kann bei einem Vergleich von zwei Datensätzen nicht vorkommen.
- kann aus der Kovarianz und den Standardabweichungen von zwei Datensätzen berechnet werden.
- liegt immer im Intervall zwischen -1 und +1.

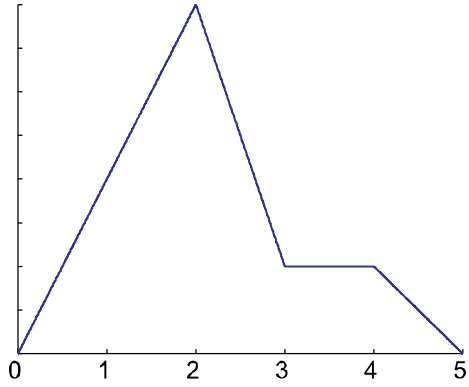
1.18 Ein bestehendes statistisches Modell, welches sowohl aleatorische als auch epistemische Unsicherheiten beinhaltet, soll mit neuen Beobachtungen aktualisiert werden. Mit Hilfe dieser neuen Beobachtungen

- kann die aleatorische Unsicherheit besser beschrieben werden.
- kann die aleatorische Unsicherheit verringert werden.
- kann die epistemische Unsicherheit verringert werden.
- verändert sich nichts an den Unsicherheiten.

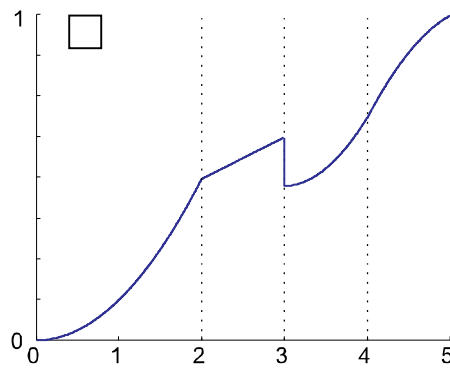
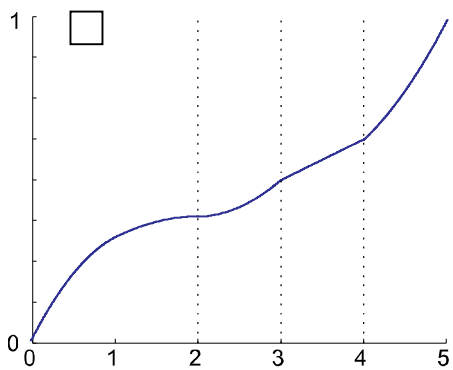
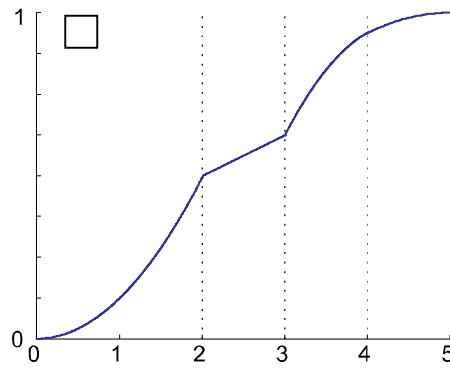
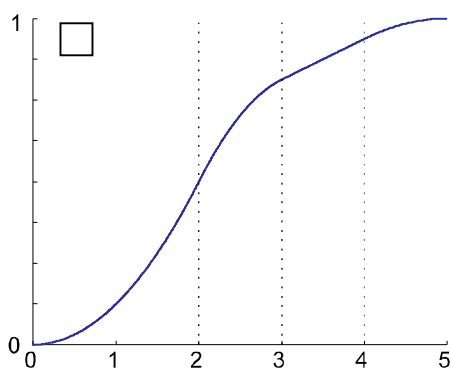
1.19 Welche der folgenden Aussagen ist/sind richtig?

- Das Integral einer Wahrscheinlichkeitsdichtefunktion von $-\infty$ bis ∞ ist immer grösser als 1.
- Das Maximum einer Wahrscheinlichkeitsverteilungsfunktion beträgt 1.
- Der Modus einer Wahrscheinlichkeitsdichtefunktion beträgt 1.
- Durch die Integration einer Wahrscheinlichkeitsverteilungsfunktion erhält man deren Dichtefunktion.

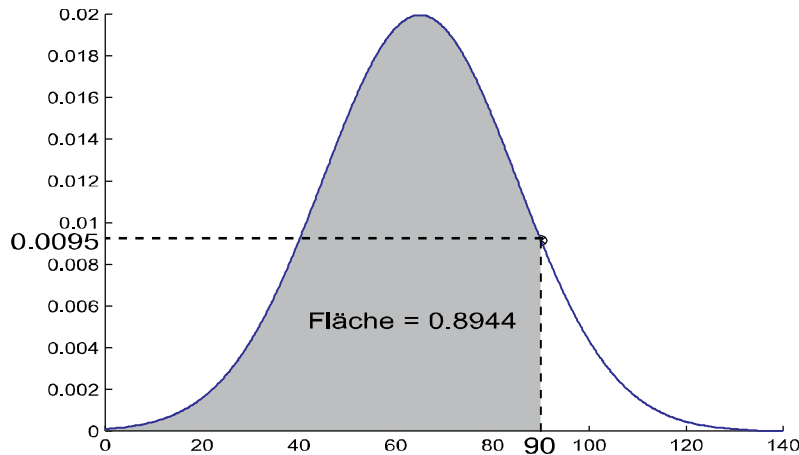
1.20 Gegeben ist folgende Wahrscheinlichkeitsdichtefunktion:



Welches ist die zu dieser Wahrscheinlichkeitsdichtefunktion passende Wahrscheinlichkeitsverteilungsfunktion?



1.21 Gegeben ist die folgende Wahrscheinlichkeitsdichtefunktion:



Wie gross ist die Wahrscheinlichkeit $P(X > 90)$, dass der Wert $x = 90$ überschritten wird?

- $P(X > 90) = 0.0095$
- $P(X > 90) = 0.8944$
- $P(X > 90) = 1 - 0.8944 = 0.1056$
- $P(X > 90) = 1 - 0.0095 = 0.9905$

1.22 Die Wahrscheinlichkeit, dass eine Zufallsvariable einen Wert in einem gegebenen Intervall annimmt, erhält man, indem man

- den Mittelwert des Intervalls in die Wahrscheinlichkeitsdichtefunktion einsetzt.
- die Wahrscheinlichkeitsdichtefunktion von der unteren bis zur oberen Intervallgrenze integriert.
- den Wert der Wahrscheinlichkeitsverteilungsfunktion der oberen Intervallgrenze von 1 abzieht.
- den Mittelwert des Intervalls in die Wahrscheinlichkeitsverteilungsfunktion einsetzt.

1.23 Welche der folgenden Umformungen ist / sind korrekt?

$$E(a + b \cdot X + c \cdot g(X)) = a + b \cdot E(X) + c \cdot E(g(X)) \quad \square$$

$$E(a + b \cdot X + c \cdot g(X)) = a + b \cdot X + c \cdot g(E(X)) \quad \square$$

$$\text{Var}(a + b \cdot X) = a + b \cdot \text{Var}(X) \quad \square$$

$$\text{Var}(a + b \cdot X) = b^2 \cdot \text{Var}(X) \quad \square$$

1.24 Die Zufallsvariable Z ist definiert durch die Linearkombination $Z = 5X + 2Y + 3$, wobei von den Zufallsvariablen X und Y jeweils die Erwartungswerte $E(X) = 2$, $E(Y) = 3$ und die Varianzen $\text{Var}(X) = 1$, $\text{Var}(Y) = 2$ bekannt sind. Die beiden Zufallsvariablen X und Y sind abhängig voneinander und die Kovarianz beträgt $C_{XY} = 0.5$.

a) Der Erwartungswert der Zufallsvariablen Z berechnet sich zu:

$$E(Z) = E(X) + E(Y) - C_{XY} = 2 + 3 - 0.5 = 4.5 \quad \square$$

$$E(Z) = 5E(X) + 2E(Y) = 5 \cdot 2 + 2 \cdot 3 = 16 \quad \square$$

$$E(Z) = 5E(X) + 2E(Y) + 3 = 3 + 5 \cdot 2 + 2 \cdot 3 = 19 \quad \square$$

$$E(Z) = 5E(X) + 2E(Y) + 2(5 \cdot 2 \cdot C_{XY}) = 5 \cdot 2 + 2 \cdot 3 + 2 \cdot 5 \cdot 2 \cdot 0.5 = 26 \quad \square$$

b) Die Varianz der Zufallsvariablen Z berechnet sich zu:

$$\text{Var}(Z) = \text{Var}(X) + \text{Var}(Y) = 1 + 2 = 3 \quad \square$$

$$\text{Var}(Z) = 5\text{Var}(X) + 2\text{Var}(Y) = 5 \cdot 1 + 2 \cdot 2 = 9 \quad \square$$

$$\text{Var}(Z) = 5\text{Var}(X) + 2\text{Var}(Y) + 2(5 \cdot 2 \cdot C_{XY}) = 5 \cdot 1 + 2 \cdot 2 + 2 \cdot 5 \cdot 2 \cdot 0.5 = 19 \quad \square$$

$$\text{Var}(Z) = 5^2\text{Var}(X) + 2^2\text{Var}(Y) + 2(5 \cdot 2 \cdot C_{XY}) = 5^2 \cdot 1 + 2^2 \cdot 2 + 2 \cdot 5 \cdot 2 \cdot 0.5 = 43 \quad \square$$

1.25 Die Kovarianz C_{XY}

- kann nur Werte im Intervall $[-1,1]$ annehmen.
- ist gleich 1, wenn die beiden Zufallsvariablen X und Y voneinander unabhängig sind.
- entspricht der Schnittmenge der Zufallsvariablen X und Y .
- beinhaltet Informationen über die lineare Abhängigkeit zwischen den Zufallsvariablen X und Y .

1.26 Die Wahrscheinlichkeitsdichtefunktion einer Summe $Z = X + Y$ von zwei voneinander unabhängigen Zufallsvariablen X und Y lässt sich berechnen, indem man

- den Erwartungswert von $X + Y$ berechnet.
- die marginalen Wahrscheinlichkeitsdichtefunktionen von X und Y addiert.
- das Faltungsintegral $f(z) = \int_{-\infty}^{\infty} f_Y(z-x)f_X(x)dx$ bildet.
- das Faltungsintegral $f(z) = \int_{-\infty}^{\infty} f_Y(y) dx \int_{-\infty}^{\infty} f_X(x)dx$ bildet.

Teil 2: Rechenaufgabe (maximal 30 Punkte)

2.1 (insgesamt 12 Punkte)

In einem Stahlwerk werden Träger hergestellt. Aus Erfahrung weiss man, dass 2% aller Träger die Anforderungen nicht erfüllen. An allen Trägern wird eine zerstörungsfreie Prüfung durchgeführt. Erfüllt ein Träger die Anforderungen nicht, wird dies von der zerstörungsfreien Prüfung zu 95% angezeigt. Erfüllt er die Anforderungen, wird dies zu 99% angezeigt.

a) (2 Punkte)

Füllen Sie die folgende Tabelle mit den bedingten Wahrscheinlichkeiten aus.

	Prüfung: nicht ausreichend I_{NA}	Prüfung: ausreichend I_A
Anforderungen erfüllt A	$P(I_{NA} A) =$	$P(I_A A) =$
Anforderungen nicht erfüllt \bar{A}	$P(I_{NA} \bar{A}) =$	$P(I_A \bar{A}) =$

b) (6 Punkte)

Berechnen Sie die Wahrscheinlichkeit, dass die Anforderungen nicht erfüllt sind, wenn die zerstörungsfreie Prüfung dies anzeigt $P(\bar{A} | I_{NA})$. Berechnen Sie ausserdem dieselbe Wahrscheinlichkeit, wenn die Prüfung anzeigt, dass die Anforderungen erfüllt sind $P(\bar{A} | I_A)$.

c) (2 Punkte)

Alternativ soll eine andere zerstörungsfreie Prüfmethode untersucht werden. Wenn der Träger die Anforderungen nicht erfüllt, wird dies von der neuen Prüfmethode zu 100% angezeigt. Sind die Anforderungen ausreichend, wird dies nur zu 98% angezeigt.

Füllen Sie die folgende Tabelle mit den bedingten Wahrscheinlichkeiten aus. (Es wird nur die neue Prüfmethode angewendet).

	Prüfung: nicht ausreichend I_{NA}	Prüfung: ausreichend I_A
Anforderungen erfüllt A	$P(I_{NA} A) =$	$P(I_A A) =$
Anforderungen nicht erfüllt \bar{A}	$P(I_{NA} \bar{A}) =$	$P(I_A \bar{A}) =$

d) (2 Punkte)

Alle als nicht ausreichend angezeigten Träger werden aussortiert. Bei welcher Prüfmethode kommen weniger Träger mit nicht ausreichenden Eigenschaften auf den Markt?

2.2 (8 Punkte)

An 10 Stahlträgern werden die tatsächlichen Festigkeiten bestimmt und mit den geschätzten Festigkeiten verglichen. In der unten gegebenen Tabelle sind die tatsächlichen Festigkeiten x_i und die geschätzten Festigkeiten y_i dargestellt (N / mm^2).

Berechnen Sie:

- die Stichprobenmittelwerte: \bar{x}, \bar{y}
- die Standardabweichungen der Stichprobe: s_X, s_Y
- die Variationskoeffizienten der Stichprobe: v_X, v_Y
- den Korrelationskoeffizienten der Stichprobe: r_{XY}

	x_i	y_i	$x_i - \bar{x}_i$	$y_i - \bar{y}_i$	$(x_i - \bar{x}_i)^2$	$(y_i - \bar{y}_i)^2$	$(x_i - \bar{x}_i)(y_i - \bar{y}_i)$
1	536	560	-29.7	-5.6	882.1	31.4	166.3
2	540	542	-25.7	-23.6	660.5	557.0	606.5
3	597	580	31.3	14.4	979.7	207.4	450.7
4	520	511	-45.7	-54.6	2088.5	2981.2	2495.2
5	570	563	4.3	-2.6	18.5	6.8	-11.2
6	567	567	1.3	1.4	1.7	2.0	1.8
7	571	561	5.3	-4.6	28.1	21.2	-24.4
8	589	601	23.3	35.4	542.9	1253.2	824.8
9	581	590	15.3	24.4	234.1	595.4	373.3
10	586	581	20.3	15.4	412.1	237.2	312.6
Σ	5657	5656			5848.1	5892.4	5195.8

2.3 (insgesamt 10 Punkte)

Die Festigkeit eines Stahlträgers in N / mm^2 wird durch eine Zufallsvariable X mit folgender Wahrscheinlichkeitsdichtefunktion dargestellt:

$$f_X(x) = \begin{cases} c & \text{für } 450 \leq x \leq 650 \\ 0 & \text{sonst} \end{cases}$$

a) (1 Punkt)

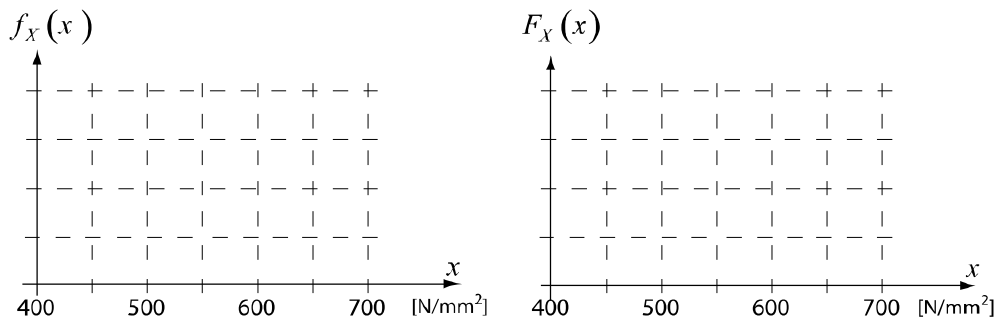
welchen Wert muss c annehmen?

b) (2 Punkte)

Geben Sie die Wahrscheinlichkeitsverteilungsfunktion $F_X(x)$ der Zufallsvariablen X an.

c) (4 Punkte)

Stellen Sie Wahrscheinlichkeitsdichtefunktion und die Verteilungsfunktion für den Bereich $400 \leq N / mm^2 \leq 700$ grafisch dar und beschriften Sie die y-Achse.



d) (2 Punkte)

Berechnen Sie den Modus, den Median und den Mittelwert.

e) (1 Punkt)

Wie gross ist die Wahrscheinlichkeit, dass die Festigkeit $\leq 500 N / mm^2$ ist?