1. VD Prüfung
Statistik und Wahrscheinlichkeitsrechnung
Herbst 2004

Prof. Dr. M.H. Faber
Prof. Dr. P. Burlando

ETH Zürich

Freitag, 1. Oktober 2004
14:00 – 16:00

Name: ....

Vorname: ....

1. VD Prüfung: Statistik und Wahrscheinlichkeitsrechnung Bau-, Umwelt- und Geomatikingenieurwissenschaften

Datum und Dauer:

Freitag, 01. Oktober 2004
Beginn: 14:00
Zeitdauer: 2 Stunden

Hilfsmittel:

Alle Unterlagen (Skripte, Bücher, andere Ausdrucke, etc.) erlaubt.
Taschenrechner erlaubt.

Administratives:

- Bitte legen Sie ihre Legi vor sich auf den Tisch.
- **Alle** Lösungsblätter müssen mit Namen, Vornamen und Studiengang versehen werden. **Nur** die zur Verfügung gestellten Blätter verwenden.
- Verwenden Sie für jede Aufgabe einen eigenen Papierbogen.
- Legen Sie am Ende der Prüfung alle Aufgaben- und Lösungsblätter in das Couvert zurück und lassen Sie dieses am Platz liegen.

Inhalt der Prüfung:

Aufgaben	Thema	Punkte
Aufgabe 1	Deskriptive Statistik -Wahrscheinlichkeitsfunktionen	2.5
Aufgabe 2	Deskriptive Statistik -Tukey-Box-Plot	10
Aufgabe 3	Bayes’scher Satz	12.5
Aufgabe 4	Wahrscheinlichkeitstheorie - Mengenlehre	12.5
Aufgabe 5	Testen von Hypothesen	10
Aufgabe 6	Chi-Quadrattest / Wahrscheinlichkeitspapier	15
Aufgabe 7	FORM - Grenzzustandsfunktion	12.5
Aufgabe 8	Bayes’sche Entscheidungstheorie	12.5
Aufgabe 9A	Wiederkehrperiode	12.5
Aufgabe 9B	Momente	
		100

Es zählt nur eine Aufgabe 9 (**entweder 9A oder 9B**), nämlich nur die Aufgabe, bei der Sie mehr Punkte erziehlen.

Hinweis:

- Die Prüfung ist vom Umfang her so konzipiert, dass alle Aufgaben 1 bis 8 und eine aus 9A und 9B gelöst werden sollen.
- Wenn Ihnen für einen Aufgabenteil ein Zwischenresultat fehlt, treffen Sie eine sinnvolle Annahme und markieren Sie diese deutlich.

Couvertinhalt:

- Allgemeine Informationen und Aufgabenstellungen (16 Seiten).
- Tabellen (3 Seiten: Normalverteilung, t-Verteilung und Chi-Quadratverteilung).
- 8 Bögen kariert.

Aufgabe 1: Deskriptive Statistik (2.5 Punkte)

1.1 Um welche Funktionsart handelt es sich in Abbildung 1? Kreuzen Sie die richtigen Lösungen an (eine richtige Lösung pro Zeile).

Tabelle 1: Funktionstypen.

☐ Wahrscheinlichkeitsdichtefunktion (= probability density function) (=pdf)	☐ Wahrscheinlichkeitsverteilungs-funktion (=cumulative distribution function) (=cdf)

☐ Stetige Funktion	☐ Diskrete Funktion

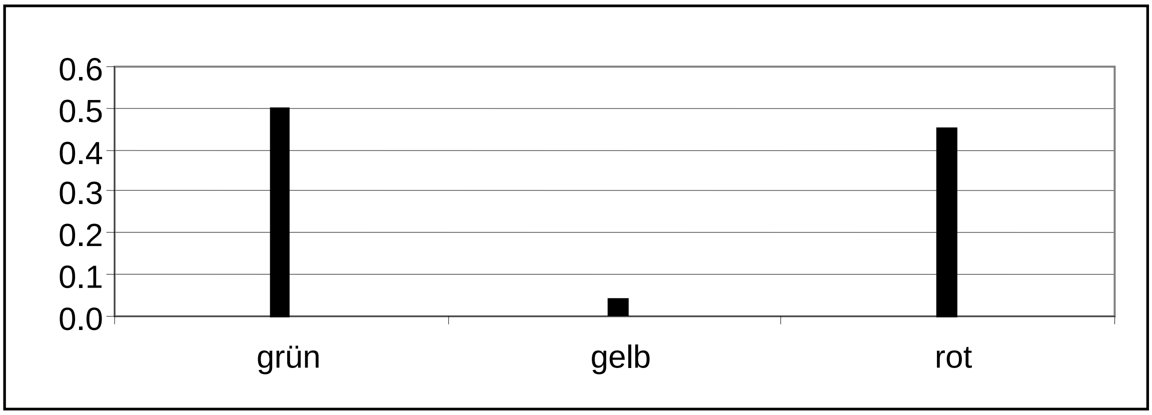


Abbildung 1: Verkehrsampel.

1.2 Um welche Funktionsart handelt es sich in Abbildung 2? Kreuzen Sie die richtigen Lösungen an (eine richtige Lösung pro Zeile).

Tabelle 2: Funktionstypen.

☐ Wahrscheinlichkeitsdichtefunktion (= probability density function) (= pdf)	☐ Wahrscheinlichkeitsverteilungs-funktion (= cumulative distribution function) (= cdf)

☐ Stetige Funktion	☐ Diskrete Funktion

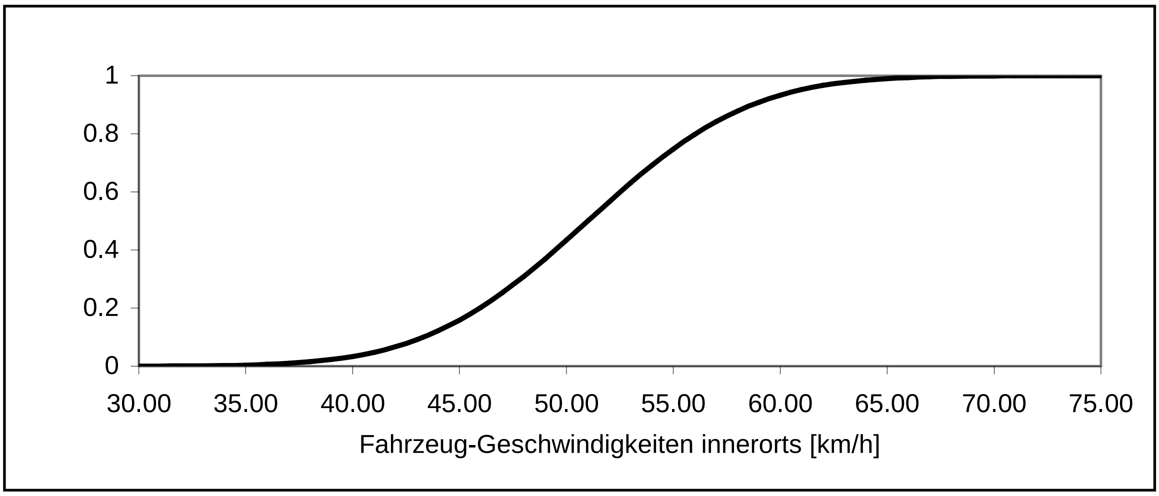


Abbildung 2: Fahrzeug-Geschwindigkeiten innerorts.

Aufgabe 2: Deskriptive Statistik - Tukey-Box-Plot (10 Punkte)

Im Juni 2004 wurden von einer Messstation im Schweizer Mittelland die folgenden Ozon-Konzentrationen gemessen (Tabelle 3):

Tabelle 3: Ozon-Konzentrationen (der Grösse nach geordnet).

Nummer	Ozon-Konzentration in [$\mu g/m^3$]
1	62
2	67
3	72
4	73
5	74
6	74
7	74
8	75
9	76
10	76
11	77
12	77
13	78
14	79
15	79
16	80
17	81
18	82
19	83
20	83
21	85
22	91
23	96
24	98
25	105
26	116
27	123
28	143
29	182
30	197

2.1 Erstellen Sie den zu diesem Datensatz dazugehörigen Tukey-Box-Plot. Berechnen Sie dazu folgendes:

Tabelle 4: Daten Tukey-Box-Plot.

	Werte in [μg/m3]
Median	
Oberes Quartil (=75%-Quantil)	
Unteres Quartil (=25%-Quantil)	
Interquartile Differenz *(r)*	
1.5-fache interquartile Differenz *(1.5r)*	
Oberer Nachbarschaftswert	
Unterer Nachbarschaftswert	
Ausreisser	

2.2 Tragen Sie diese Werte im untenstehenden Diagramm in Form eines Tukey-Box-Plots ein (Abbildung 3):

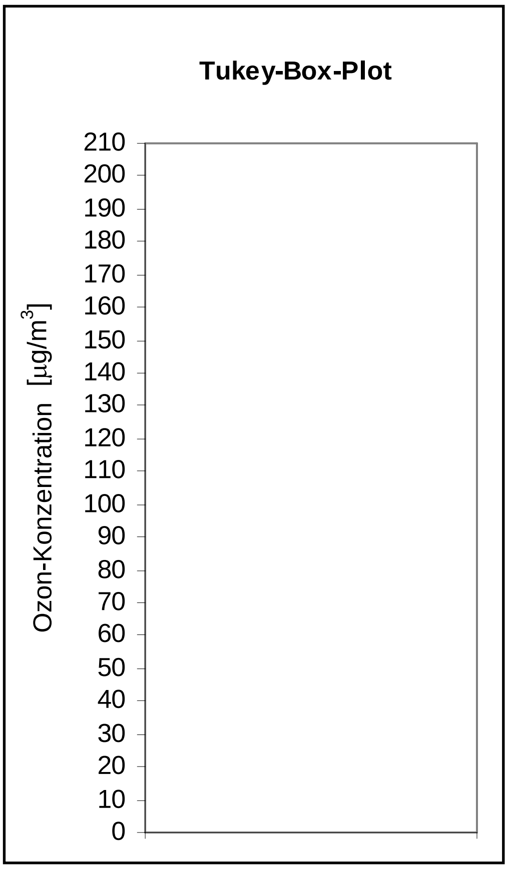


Abbildung 3: Tukey-Box-Plot.

2.3 Kreuzen Sie die richtigen Antworten bezüglich der Verteilung der Ozon-Konzentrationen (Tabelle 5) an. Es gibt eine richtige Antwort pro Zeile.

Tabelle 5: Fragen zu Tukey-Box-Plot und Standardabweichung.

□ Die Verteilung ist linksschief / rechtssteil	□ Die Verteilung ist rechtsschief / linkssteil
□ Der Mittelwert ist grösser als der Median	□ Der Mittelwert ist kleiner als der Median
□ Der Grenzwert gemäss Luftreinhalte-verordnung von 120 $\mu g/m^3$ wird von weniger als von 25% der Werte überschritten.	□ Der Grenzwert gemäss Luftreinhalte-verordnung von 120 $\mu g/m^3$ wird von mehr als von 25% der Werte überschritten.
Es sind die Standardabweichungen s_{biased} und $s_{unbiased}$ ermittelt worden. Dabei wurden zwei Werte (35.4 $\mu g/m^3$ und 36.0 $\mu g/m^3$) berechnet. Welcher Wert entspricht welcher Standardabweichung?	
□ s_{biased} = 35.4 $\mu g/m^3$ $s_{unbiased}$ 36.0 $\mu g/m^3$	□ s_{biased} = 36.0 $\mu g/m^3$ $s_{unbiased}$ 35.4 $\mu g/m^3$

Aufgabe 3: Bayes'scher Satz (12.5 Punkte)

Für die Geologie-Prüfung muss sich ein Student Kenntnisse über Mineralien und Gesteine aneignen. Zu diesem Zweck hat er sich in den Geologie-Keller der ETH begeben.

Bei einem Stein kann er sich nicht entscheiden: Ist es Calzit oder Gips?
Er schätzt die Situation a priori wie folgt ein:
- Wahrscheinlichkeit, dass der Stein ein Calzit ist, beträgt P(C)=0.4.
- Wahrscheinlichkeit, dass der Stein ein Gips ist, beträgt P(G)=0.6.

Um die Sachlage zu klären, führt der Student einen Test mit Salzsäure durch. Denn theoretisch gilt folgendes:
- Salzsäure auf Calzit führt zu einem Brausen.
- Salzsäure auf Gips führt nicht zu einem Brausen.

Aufgrund von Heterogenitäten des Gesteins (z.B. Einschlüsse anderer Mineralien), können beim Salzsäuretest auch Fehler auftreten. Für den Salzsäuretest gelten die folgenden Parameter:

Tabelle 6: Eigenschaften des Salzsäuretests.

		Ergebnisse des Salzsäuretests	
		Brausen: Indikation auf Calzit (I=C)	Kein Brausen: Indikation auf Gips (I=G)
Tatsächlicher Zustand	Calzit (C)	P(I=C|C)=0.99	P(I=G|C)=0.01
	Gips (G)	P(I=C|G)=0.10	P(I=G|G)=0.90

Der Student trägt Salzsäure auf den fraglichen Stein auf. Resultat: Es braust!

3.1 Wie gross ist die Wahrscheinlichkeit, dass es sich (trotz Brausen beim Salzsäuretest) um einen Gips handelt?

3.2 Im vorliegenden Fall ist der Student davon ausgegangen, dass es sich mit einer Wahrscheinlichkeit von P(G)=0.6 um einen Gips handelt.

Wie gross müsste seine a-priori Schätzung P(G), dass das vorliegende Gestein ein Gips ist, mindestens sein, dass er sich trotz Brausen beim Salzsäuretest am Schluss für den Gips entscheidet? (Tipp: Die Entscheidungsgrenze liegt bei einer a-posteriori Wahrscheinlichkeit von P(G|I=C)=0.5).

Kommentieren Sie das Resultat: Würde der Student bei einer solchen a-priori Schätzung überhaupt einen Salzsäuretest durchführen? Warum?

3.3 Beschreiben Sie die Formel ***P(I=C|G)*** kurz in eigenen Worten.

3.4 Beschreiben Sie die Formel ***P(G|I=C)*** kurz in eigenen Worten.

Aufgabe 4: Wahrscheinlichkeitstheorie - Mengenlehre (12.5 Punkte)

Ein mittelgrosses Unternehmen im bernischen Emmental verwendet für seine Produktion grössere Mengen an Propan, welches in einem Tank auf dem Betriebsgelände gelagert wird (Abbildung 4).

Abbildung 4: Propantank.

Propan ist ein hochentzündliches Gas. Deshalb muss zu Handen des Kantons Bern das Risiko des Propantanks berechnet werden. Zu diesem Zweck ist ein Fehlerbaum erstellt worden, mit dem die Wahrscheinlichkeit eines Feuerballs (=BLEVE) berechnet werden soll (Abbildung 5).

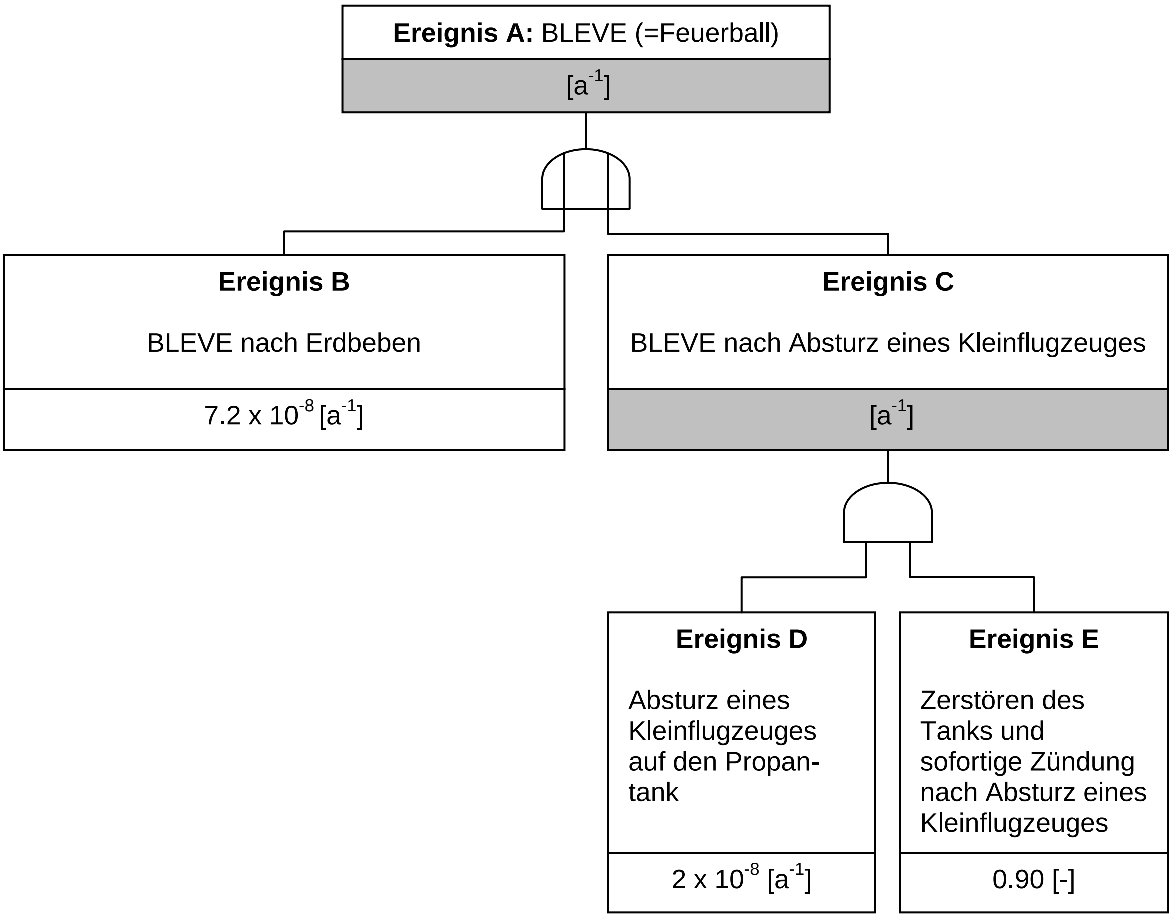
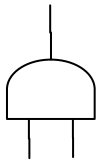
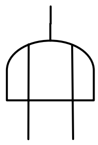


Abbildung 5: Wahrscheinlichkeit eines BLEVEs. (Die Einheit [a^{-1}] bedeutet pro Jahr).

Hinweis:



Dieses Symbol stellt eine **UND**-Verknüpfung zweier Ereignisse dar.



Dieses Symbol stellt eine **ODER**-Verknüpfung zweier Ereignisse dar.

Abbildung 6: Symbole der UND- sowie der ODER-Verknüfung zweier Ereignisse.

4.1 Berechnen Sie die Wahrscheinlichkeit des Ereignisses C (=unteres graues Feld in der Abbildung 5).

4.2 Es wird angenommen, dass die Ereignisse B und C „statistisch unabhängig“ sind. Diskutieren Sie anhand dieses Beispiels, was statistisch unabhängig bedeutet.

4.3 Berechnen Sie die Wahrscheinlichkeit des Ereignisses A (= oberes graues Feld in der Abbildung 5). Falls Sie die Aufgabe 4.1 nicht gelöst haben, nehmen Sie für P[C] den Wert $1 \times 10^{-8}\ a^{-1}$ an.

4.4 Falls bei einem BLEVE ein Sachschaden von 30 Mio. Franken entsteht, wie gross ist das Risiko in [Fr./a] eines BLEVE? Falls Sie die Aufgabe 4.3 nicht gelöst haben, nehmen Sie für P[A] den Wert $8 \times 10^{-8}\ a^{-1}$ an.

Aufgabe 5: Testen von Hypothesen (10 Punkte)

Aus langjährigen von Personen durchgeführten Verkehrsbeobachtungen ist bekannt, dass die für die Parkplatzsuche benötigte Zeit in Zürich normalverteilt angenommen werden kann mit einem Mittelwert von 30 Minuten.

Da von Personen durchgeführte Verkehrsbeobachtungen teuer sind, wird der Einsatz von automatisierten Verkehrsüberwachungssystemen in Erwägung gezogen. Das Verhalten des neuen Systems ist unbekannt, weshalb man einen Testlauf durchführt.

5.1 In einem solchen Testlauf wird die Parkplatzsuchzeit von 30 Fahrzeugen erfasst. Der Mittelwert der Parkplatzsuchzeiten beträgt für diese Stichprobe 32 Minuten, die Standardabweichung (unbiased) 4 Minuten. Testen Sie auf ein Signifikanzniveau von 5 % die Hypothese, dass das neue automatisierte Verkehrsüberwachungssystem den gleichen Mittelwert liefert wie die von Personen durchgeführten Verkehrsbeobachtungen.

5.2 Nehmen Sie an, dass von der Grundgesamtheit der Parkplatzsuchzeit der Mittelwert 30 Minuten und die Standardabweichung 4 Minuten betragen. Wieviele Beobachtungen von parkplatzsuchenden Fahrzeugen sind erforderlich, damit der Mittelwert der Stichprobe, auf ein Signifikanzniveau von 5%, innerhalb von ±3 Minuten um den Mittelwert der Grundgesamtheit liegt.

5.3 Hypothesen können angenommen oder verworfen werden. Abhängig vom Signifikanzniveau, auf dem die Tests durchgeführt werden, ist die Wahrscheinlichkeit grösser oder kleiner, dass falsche Schlussfolgerungen gezogen werden können. Bei falschen Schlussfolgerungen wird nach Fehler 1. Art und Fehler 2. Art unterschieden. Tragen Sie die entsprechenden Buchstaben A, B oder C in die geeigneten Felder der Tabelle 7 ein:

A: Korrekte Entscheidung
B: Fehler 1. Art
C: Fehler 2. Art

Tabelle 7: Auswirkungen von Annehmen oder Verwerfen von Hypothesen H_0.

		Entscheidung	
		H_0 akzeptieren	H_0 verwerfen
Tatsächliche Situation	H_0 ist richtig		
	H_0 ist falsch		

Aufgabe 6: Chi-Quadrat-Test, Wahrscheinlichkeitspapier (15 Punkte)

Die Permeabilität von Beton ist von hoher Bedeutung für die Lebensdauer von Betonbauwerken. Deshalb wurden in einem neugebauten Tunnel an 24 Stellen Permeabilitätsmessungen durchgeführt. Für jede Stelle wurde der Permeabilitätskoeffizient kT bestimmt.

Langjährige Erfahrung: Aus früheren Messungen an ähnlichen Bauwerken ist bekannt, dass die Permeabilität normalverteilt ist, mit einem Mittelwert von 0.15×10^{-16} m^2.

Stichprobe (neuer Tunnel): Die Standardabweichung der Stichprobe der Messungen im neuen Tunnel beträgt 0.05×10^{-16} m^2.

6.1 Führen Sie einen Chi-Quadrattest durch. Testen Sie auf dem 5%-Signifikanzniveau, ob die gemessenen Daten im Tunnel (siehe Tabelle 8) durch eine Normalverteilung mit den oben angegebenen Parametern beschrieben werden kann.

Tabelle 8: Chi-Quadrattest für Permeabilitätskoeffizientsdaten in einem neuen Tunnel.

Klassen	Permeabilitäts-koeffizient in [10^{-16} m^2]		Anzahl beobachtete Messungen pro Klasse (n_i)	Theoretisches Wahrscheinlichkeitsmodell		Testgrösse
	Untere Klassen-grenze	Obere Klassen-grenze		Theoretische Wahrschein-lichkeit pro Klasse (Δp_i)	Theoretische Anzahl Werte pro Klasse (r_i)	
Klasse 1	0	0.1	6			
Klasse 2	0.1	0.2	13			
Klasse 3	0.2	∞	5			

6.2 Es ist bereits ein Wahrscheinlichkeitspapier erstellt worden (siehe Abbildung 7), um zu überprüfen, ob die im neuen Tunnel gemessenen Permeabilitätsdaten durch eine Normalverteilung beschrieben werden können. Aus Abbildung 7 ist ersichtlich, dass dies der Fall ist, denn die Daten liegen ungefähr auf einer Gerade.

Schätzen Sie mit Hilfe des Wahrscheinlichkeitspapiers die Verteilungsparameter.

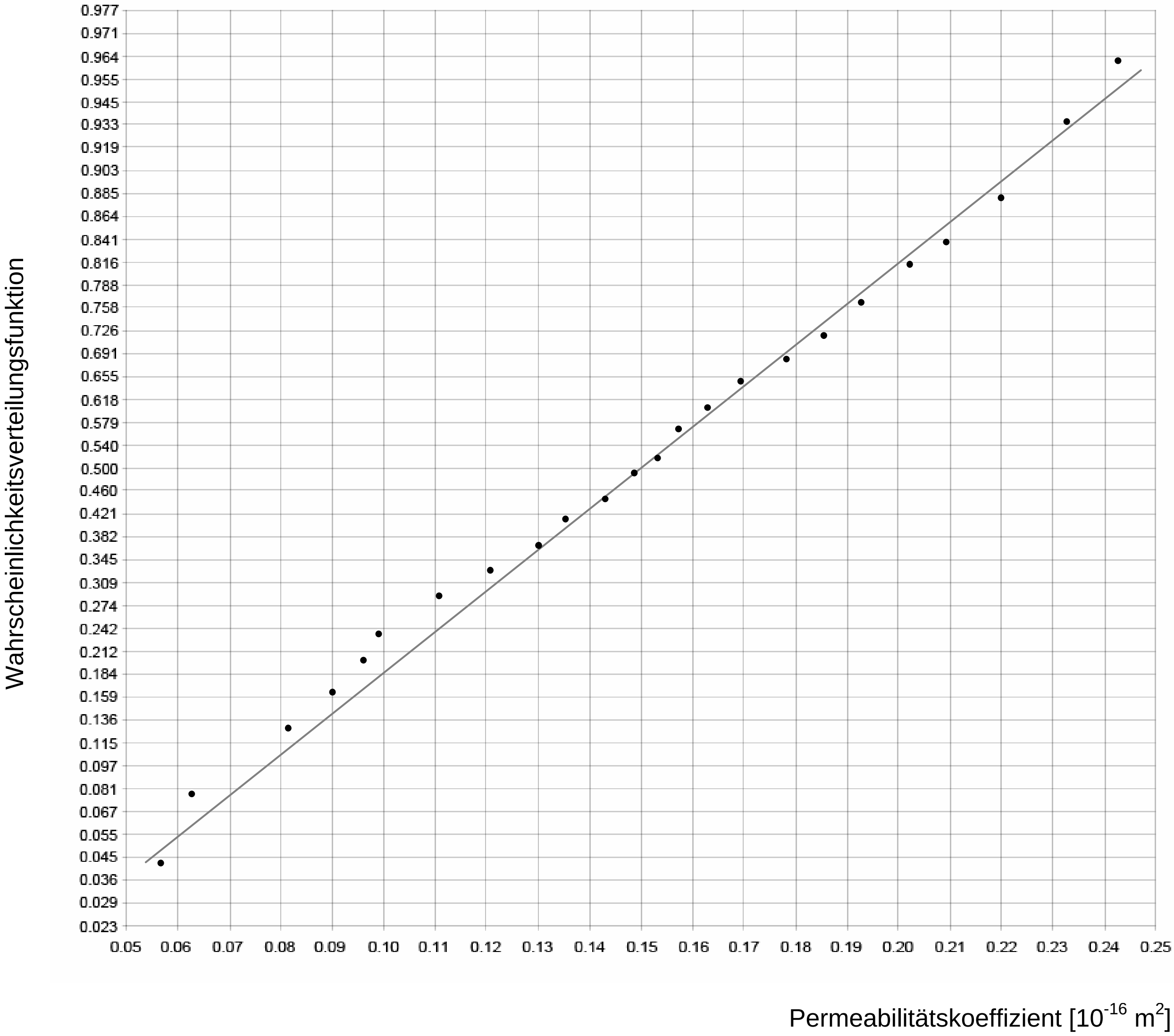


Abbildung 7: Wahrscheinlichkeitspapier für die 24 Messwerte des Permeabilitätskoeffizienten im neuen Tunnel.

Aufgabe 7: FORM - Grenzzustandfunktion (12.5 Punkte)

In der Risikoabschätzung für den Propantank (gemäss Aufgabe 4) fehlt den Behörden des Kantons Bern die Abschätzung der Gefahr durch ein von der Strasse abkommendes Auto, das zwischen den Produktionsgebäuden hindurch in den Propantank prallen könnte.

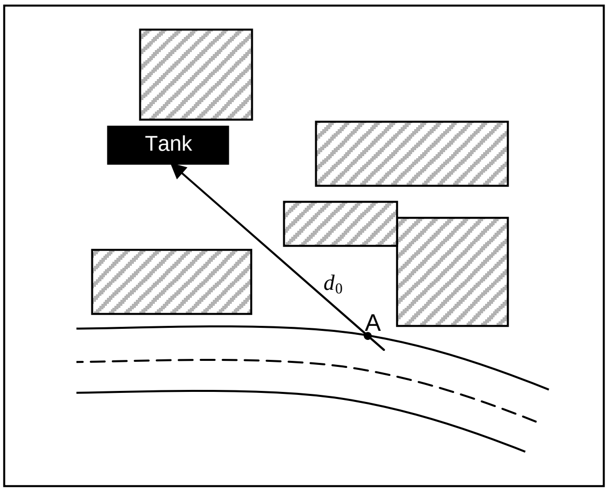


Abbildung 8: Beschädigung des Tanks durch Abkommen eines Autos von der Strasse.

Folgende Informationen sind gegeben:

- Es wird angenommen, dass die gefährdungsrelevanten Autos in Punkt A abkommen, gleich zu bremsen beginnen und sich auf der dargestellten Gerade in Richtung Tank bewegen.
- Distanz d_0 von der Strasse zum Tank: 38 m.
- Die Geschwindigkeit v der Autos sei normalverteilt mit μ_v = 14.4 m/s (entspricht 52 km/h) und σ_v = 1.7 m/s (entspricht 6 km/h). Tipp: Immer die Masseinheiten Meter und Sekunden verwenden.
- Die Bremsbeschleunigung a der Autos sei normalverteilt mit μ_a = 4.5 m/s^2 und σ_a = 1.3 m/s^2.
- Die Masse der abkommenden Autos sei m (siehe Abbildung 9).

Mit einer vereinfachten Energiebetrachtung lässt sich die Distanz d, bis in welche ein von der Strasse abkommendes Auto erreicht, abschätzen:

$$E_{kin} = E_{brems}$$
$$\frac{1}{2} m\, v^2 = m\, d\, a$$
$$d = \frac{v^2}{2a}$$

Abbildung 9: Von einem abkommenden Auto erreichte Distanz d .

7.1 Bis in welche Distanz d gelangt (gemäss der Formel in Abbildung 9) ein Auto mit Geschwindigkeit v=14.4 m/s und Bremsbeschleunigung a=4.5 m/s²?

7.2 Stellen Sie die Gleichung auf der Grenzzustandsfunktion g als Funktion der standardisierten stochastischen Variablen z_v und z_a auf.

7.3 Führen Sie die Iterationsschritte 2 und 3 durch in Tabelle 10 durch (leere Felder) und bestimmen Sie den Sicherheitsindex β. Sämtliche Iterationsformeln sind bereits berechnet worden (Tabelle 9).

Tabelle 9: Iterationformeln.

Iterationsformel für β: $\beta_{neu} = \dfrac{-134.64}{98.8 \cdot \alpha_a - 2.89 \cdot \beta_{alt} \cdot \alpha_v^{\ 2} - 48.96 \cdot \alpha_v}$	
Iterationsformeln für α_v', α_a'.(nicht normiert): $\alpha_v' = (-5.78 \cdot \beta \cdot \alpha_v - 48.96)$	$\alpha_a' = 98.8$
Iterationsformel für k: $k = \sqrt{(\alpha_v')^2 + (\alpha_a')^2}$	
Iterationsformeln für α_v, α_a.(normiert): $\alpha_v = -\dfrac{1}{k}\alpha_v'$	$\alpha_a = -\dfrac{1}{k}\alpha_a'$

Tabelle 10: Iterationsverfahren zur Bestimmung des Grenzpunktes β.

Iterationsschritt	β	k	α's normiert	
			α_v	α_a
Startwerte	4	1	0	-1
1. Schritt	1.36	110.27	0.44	-0.90
2. Schritt				
3. Schritt				

7.4 Wie gross ist die Wahrscheinlichkeit, dass von der Strasse abkommende Fahrzeuge in den Tank prallen?

Falls Sie Aufgabe 7.3 nicht berechnet haben, verwenden Sie für den Sicherheitsindex β einen Wert von 1.3.

Aufgabe 8: Bayes'sche Entscheidungstheorie (12.5 Punkte)

Es sollen die kostenoptimalen Optionen für die Umnutzung eines alten Bauwerks zu einem Bürogebäude untersucht werden. Dabei wird angenommen, dass die Druckfestigkeit des Betons der Deckenplatten massgebend für die Versagenswahrscheinlichkeit des Bauwerks ist. Die Druckfestigkeit der Deckenplatten wird in drei Klassen eingestuft.

Klasse 1: θ_1: 15 MPa
Klasse 2: θ_2: 20 MPa
Klasse 3: θ_3: 25 MPa

Die tatsächlich vorhandene Druckfestigkeit der Deckenplatten des alten Bauwerks ist unbekannt. Mit folgenden a-priori Wahrscheinlichkeiten ist die entsprechende Klasse vorhanden:

Klasse 1: $P'(\theta_1) = 0.3$
Klasse 2: $P'(\theta_2) = 0.45$
Klasse 3 $P'(\theta_3) = 0.25$

Als mögliche Optionen werden identifiziert:

a_1: unbeschränkte Nutzung nach Einbau neuer Deckenplatten mit einer Betondruckfestigkeit von 25 MPa.
a_2: Nutzung auf leichtes Mobiliar beschränkt (ohne schwere Gegenstände, z.B. Ordnerschränke).
a_3: Nutzung unbeschränkt.

Die Versagenswahrscheinlichkeit der Deckenplatten bei unbeschränkter Nutzung in Abhängigkeit von den Druckfestigkeitsklassen (θ_i) beträgt:

$$P(V \mid \theta_1) = 0.04$$
$$P(V \mid \theta_2) = 0.03$$
$$P(V \mid \theta_3) = 0.02$$

Die Versagenswahrscheinlichkeiten der Deckenplatten in Abhängigkeit von den Druckfestigkeitsklassen (θ_i) reduzieren sich bei der beschränkten Nutzung jeweils um 1%.

Die Kosten für den Einbau neuer Deckenplatten betragen 50'000 CHF. Die Kosten für die beschränkte Nutzung werden auf 10'000 CHF geschätzt. Die Versagenskosten betragen 3'000'000 CHF.

Erstellen Sie einen a-priori Ereignis-/Entscheidungsbaum und bestimmen Sie die optimale Option.

Aufgabe 9A: Wiederkehrperiode (12.5 Punkte)

Nehmen Sie an, dass eine Einwirkung mit der Extremalverteilung Typ I (Gumbel max) beschrieben wird. Die Verteilungsfunktion ist gegeben mit:

$$F_{X,T}^{max}(x) = exp\left(-exp\left(-\alpha\left(x-u\right)\right)\right)$$

Für die Bemessung wird der charakteristische Wert x_c für die jährliche Überschreitungswahrscheinlichkeit p benötigt.

Zeigen Sie, dass dieser charakteristische Wert x_c näherungsweise mit folgender Formel berechnet werden kann.

$$x_c \approx \frac{1}{\alpha} ln\left(T_R\right) + u$$

x_c : charakteristischer Wert für die jährliche Überschreitungswahrscheinlichkeit p
u : Parameter der Gumbelverteilung
α : Parameter der Gumbelverteilung
T_R : Wiederkehrperiode

Hinweis: Verwenden Sie als Näherung für $ln(1-p)$ die Taylorreihen-entwicklung erster Ordnung im Entwicklungspunkt $p=0$.

Aufgabe 9B: Momente (12.5 Punkte)

Zeigen Sie, dass gilt:

$$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y] + 2ab Cov[X,Y]$$

Dabei dürfen Sie folgendes als vorausgesetzt betrachten:

$$\int_{-\infty}^{\infty} f_X(x) = 1 \qquad (1)$$

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx \qquad (2)$$

$$Var[X] = E[X^2] - E[X]^2 \qquad (3)$$

$$E[X+Y] = E[X] + E[Y] \qquad (4)$$

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x,y) dx dy \qquad (5)$$

$$Cov[X,Y] = E[XY] - E[X]E[Y] \qquad (6)$$

Beschreiben Sie bei jedem Rechenschritt, welche Voraussetzung verwendet wurde, indem Sie die entsprechende Gleichungsnummer angeben.