Basisprüfung B. Sc.

Statistik und Wahrscheinlichkeitsrechnung

FS 2009

(Als Tafelübung im FS 2011, Aufgabe 1c leicht abgeändert)

Musterlösung

Inhalt der Prüfung:

Inhalt	**Aufgaben**	**Seite**	**Punkte**
Aufgabe 1	Beschreibende Statistik, Signifikanz-Test	2-3	20
Aufgabe 2	Modellbildung, Modellevaluation	4.5	30
Aufgabe 3	Versagenswahrscheinlichkeit, Grenzzustandsfunktion	6	20
Aufgabe 4	Entscheidungsanalyse	7-8	25
Aufgabe 5	Extremwerte, Zufallsprozesse	9	25
Anhang	Tabellen	10-12	-

Aufgabe 1: Beschreibende Statistik und Signifikanz-Test (20 Punkte)

Daten zu Lawinenunfällen:

Tabelle 1.1: *Lawinenunfallstatistik*

Jahr	Anzahl von Lawinen erfasster Personen	Anzahl von Lawinen getöteter Personen
1994	118	21
1995	108	20
1996	140	18
1997	171	24
1998	83	13
1999	215	35

a) *(4 Punkte)* **Streudiagramm**

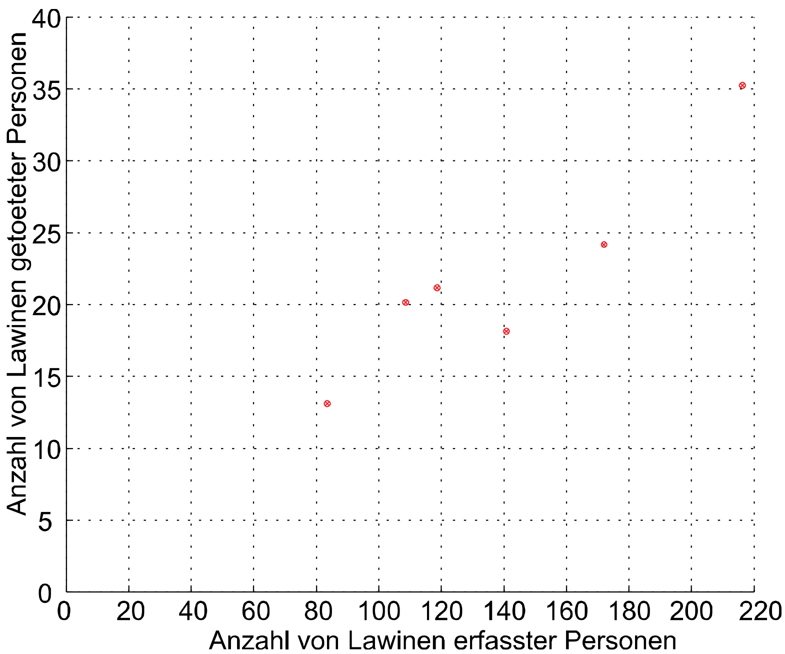


Abbildung 1.1: *Streudiagramm.*

Die Korrelation ist positiv .

b) *(8 Punkte)* **Quantilplot**

i	Anzahl von Lawinen getöteter Personen	Quantilindex
1	13	i/(n+1)=1/7=0.1429
2	18	0.2857
3	20	0.4286
4	21	0.5714
5	24	0.7143
6	35	0.8571

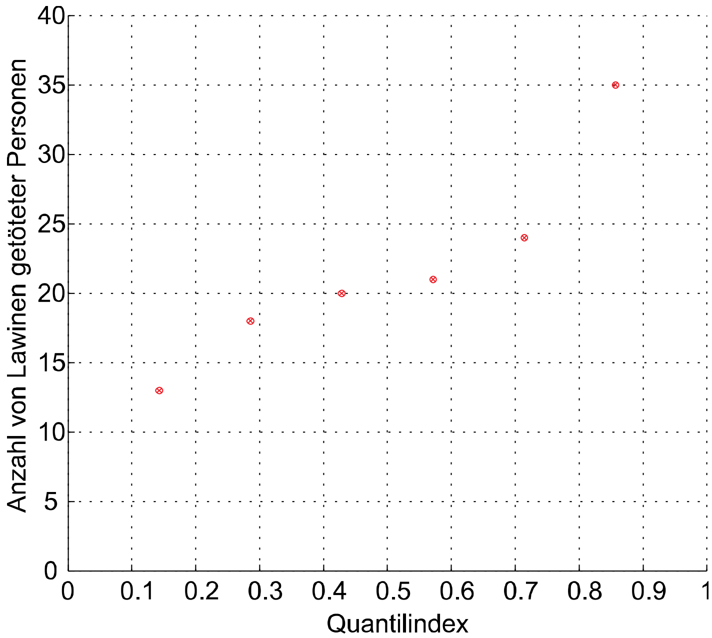


Abbildung 2.1: *Quantilplot der Anzahl von Lawinen getöteter Personen.*

Mittelwert, Median und 0.75-Quantil

Mittelwert: $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i = \frac{131}{6} = \underline{\underline{21.83}}$

Median: $\frac{x_3^o + x_4^o}{2} = \frac{20+21}{2} = \underline{\underline{20.5}}$

0.75-Quantil:

Bestimmung des Rangs: $i = v(n+1) = 0.75 \cdot 7 = 5.25$

Interpolation: $x_i^o = x_5^o + 0.25(x_6^o - x_5^o) = 24 + 0.25 \cdot (35-24) = \underline{\underline{26.75}}$

c) *(8 Punkte)* **Hypothesentest**

Nullhypothese:
$H_0 : \mu_{1994-1999} = \mu_{1980-1993} = 87$

Operative Regel:
Akzeptanz der Nullhypothese wenn $\mu_X - \Delta \leq \overline{x} \leq \mu_X + \Delta$

Signifikanzniveau:
$\alpha = 0.05$

Berechnung des Akzeptanzkriteriums:

$$P[\mu_X - \Delta \leq \overline{X} \leq \mu_X + \Delta] = 1 - \alpha$$

$$\Delta = k_{\alpha/2} \frac{\sigma_X}{\sqrt{n}} = k_{0.975} \frac{\sigma_X}{\sqrt{n}} = 1.96 \cdot \frac{48}{\sqrt{6}} = 38.41$$

$$P[87 - 38.41 \leq \overline{X} \leq 87 + 38.41] = 0.95$$

$$\Downarrow$$

$$\underline{\underline{P[48.59 \leq \overline{X} \leq 125.41] = 0.95}}$$

Berechnung des Stichprobenmittelwertes (Spalte 2 in Tabelle 1.1):

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} \hat{x}_i = \frac{835}{6} = 139.17$$

<u>Rückschluss:</u>
Da der Stichprobenmittelwert nicht im Intervall liegt, muss die Nullhypothese auf einem Signifikanzniveau von $\alpha = 0.05$ verworfen werden.

Aufgabe 2: Modellbildung, Modellevaluation (30 Punkte)

a) *(8 Punkte)* **Wahrscheinlichkeitspapier (Normalverteilung)**

Linearisierung:

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\Phi^{-1}\left(F_X(x)\right) = \frac{x-\mu}{\sigma}$$

i	Anzahl von Lawinen getöteter Personen	$F^o(x)=i/(n+1)$	$\Phi^{-1}(F^o(x))$
1	13	0.1429	-1.07
2	18	0.2857	-0.57
3	20	0.4286	-0.18
4	21	0.5714	0.18
5	24	0.7143	0.57
6	35	0.8571	1.07

Tipp zum Auslesen von Werten aus der Tabelle:

$$\Phi^{-1}(0.1429) = -\Phi^{-1}(1-0.1429) = -\Phi^{-1}(0.8571) \approx -1.07$$

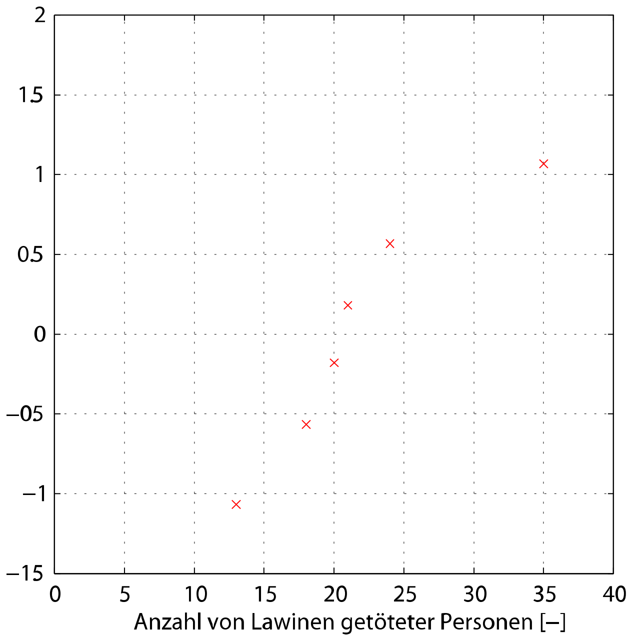


Abbildung 2.1: *Wahrscheinlichkeitspapier für eine Normalverteilung.*

<u>Antwortsatz</u>: Die Normalverteilung ist einigermassen dazu geeignet, die Daten zu repräsentieren, aufgrund der geringen Datenmenge ist allerdings keine eindeutige Aussage möglich.

b) *(8 Punkte)* **Methode der Momente**

Normalverteilung:

$$\mu = \bar{x} = \underline{\underline{21.83}}$$

$$s_X^2 = m_2 - m_1^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i^2 - \bar{x}^2 = 522.5 - 21.83^2 = 45.8$$

$$\sigma = s_X = \sqrt{45.8} = \underline{\underline{6.77}}$$

Lognormalverteilung:

$$\mu = \exp(\lambda + \frac{\zeta^2}{2})$$

$$\sigma = \exp(\lambda + \frac{\zeta^2}{2})\sqrt{\exp(\zeta^2) - 1} = \mu\sqrt{\exp(\zeta^2) - 1}$$

$$\rightarrow \left(\frac{\sigma}{\mu}\right)^2 + 1 = \exp(\zeta^2) \quad \rightarrow \quad \ln\left[\left(\frac{\sigma}{\mu}\right)^2 + 1\right] = \zeta^2$$

$$\zeta^2 = \ln\left[\left(\frac{6.77}{21.83}\right)^2 + 1\right] = 0.09183$$

$$\underline{\underline{\zeta = 0.303}}$$

$$\ln\mu = \lambda + \frac{\zeta^2}{2}$$

$$\lambda = \ln\mu - \frac{\zeta^2}{2} = \ln 21.83 - \frac{0.09183}{2}$$

$$\underline{\underline{\lambda = 3.03737}}$$

c) *(10 Punkte)* **Stichproben-Log-Likelihood**

Für die Lognormalverteilung.:

$$l(\boldsymbol{\theta} \mid \hat{\mathbf{x}}) = \ln L(\boldsymbol{\theta} \mid \hat{\mathbf{x}}) = \ln \prod_{i=1}^{n} f_X(\hat{x}_i \mid \boldsymbol{\theta}) = \sum_{i=1}^{n} \ln f_X(\hat{x}_i \mid \boldsymbol{\theta})$$

$$= \sum_{i=1}^{n} \ln \left(\underbrace{\frac{1}{\hat{x}_i \zeta \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln(\hat{x}_i) - \lambda}{\zeta} \right)^2 \right)}_{f_X(\hat{x}_i \mid \boldsymbol{\theta})} \right) = \sum_{i=1}^{n} \ln \left(\frac{1}{\hat{x}_i 0.4 \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln(\hat{x}_i) - 3}{0.4} \right)^2 \right) \right)$$

Parameter: $\lambda = 3.0$ *und* $\zeta = 0.4$

x_i	$f(x_i)$	$\ln(f(x_i))$
21	0.0472	-3.05336
20	0.049865	-2.99844
18	0.053366	-2.93058
24	0.037637	-3.27977
13	0.042466	-3.15906
35	0.01087	-4.52178
	Summe	**-19.943**

Rückschluss:

Da die Log-Likelihood der Lognormalverteilung (-19.943) grösser ist als die der Normalverteilung (-19.987), repräsentiert die Lognormalverteilung die Beobachtungen besser.

Alternativ: Vereinfachung der Log-Likelihood-Funktion:

$$l(\boldsymbol{\theta} \mid \hat{\mathbf{x}}) = \sum_{i=1}^{n} \ln \left(\frac{1}{\hat{x}_i \zeta \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln(\hat{x}_i) - \lambda}{\zeta} \right)^2 \right) \right)$$

$$= \sum_{i=1}^{n} -\ln(\hat{x}_i) - \ln\left(\zeta\sqrt{2\pi}\right) - \frac{1}{2}\left(\frac{\ln(\hat{x}_i) - \lambda}{\zeta} \right)^2$$

$$= -n \ln\left(\zeta\sqrt{2\pi}\right) - \sum_{i=1}^{n} \ln(\hat{x}_i) - \frac{1}{2\zeta^2} \sum_{i=1}^{n} \left(\ln(\hat{x}_i) - \lambda \right)^2$$

$$= -6 \cdot \ln\left(0.4\sqrt{2\pi}\right) - 18.229 - \frac{1}{2 \cdot 0.4^2} \cdot 0.5434$$

$$= \underline{\underline{-19.943}}$$

d) *(4 Punkte)* **Kolmogorov-Smirnov-Test**

Normalverteilung mit Parametern $\mu = 23$ und $\sigma^2 = 25$
Signifikanzniveau $\alpha = 0.05$

Tabelle 2.1: *Angaben zur Hilfestellung für die Durchführung eines Kolmogorv-Smirnov-Tests.*

Rang i	Beobachtung x	$F_X(x)$	$\frac{i}{n}$	$\left\lvert \frac{i}{n} - F_X(x) \right\rvert$
1	13	0.0228	0.1667	0.1439
2	18	0.1587	0.3333	0.1746
3	20	0.2743	0.5000	0.2257
4	21	0.3446	0.6667	**0.3221**
5	24	0.5793	0.8333	0.2540
6	35	0.9918	1.0000	0.0082

maximale Abweichung aus den Beobachtungen: $\varepsilon_{\max} = 0.3221$
Auslesen des kritischen Wertes aus Tabelle 4: n=6; α=0.05→ $c = 0.519$

Rückschluss:
Da $\varepsilon_{\max} = 0.3221$ kleiner ist als $c = 0.519$, muss die Verteilung auf einem Signifikanzniveau von $\alpha = 0.05$ nicht verworfen werden.

Aufgabe 3: Versagenswahrscheinlichkeit und Grenzzustandsfunktion (20 Punkte)

a) *(10 Punkte)* **Normalverteilte Zufallsvariablen**

Grenzzustandsfunktion für das Loslösen eines Schneebretts: $M = (F_T + F_C + F_F + F_S) - T$

$$\mu_M = (\mu_{F_T} + \mu_{F_C} + \mu_{F_F} + \mu_{F_S}) - \mu_T$$
$$= 315 + 520 + 420 + 1900 - 2000 = 3155 - 2000 = \underline{1155}$$

$$\sigma_M = \sqrt{(\sigma_{F_T})^2 + (\sigma_{F_C})^2 + (\sigma_{F_F})^2 + (\sigma_{F_S})^2 + (\sigma_T)^2} = \sqrt{(30)^2 + (50)^2 + (40)^2 + (190)^2 + (200)^2}$$
$$= \sqrt{900 + 2500 + 1600 + 36100 + 40000} = \sqrt{81100} = \underline{285}$$

$$\beta = \frac{\mu_M}{\sigma_M} = \frac{1155}{285} = \underline{\underline{4.056}}$$

$$p_f = \Phi(-\beta) = \Phi(-4.056) = 1 - \Phi(4.06) = 1 - 0.99997546 = \underline{\underline{2.454 \cdot 10^{-05}}}$$

b) *(10 Punkte)* **Summe von gleichverteilten Zufallsvariablen**

- Kosten: Summe K_{ges} aller 50 gleichverteilten, unabhängigen Zufallsvariablen K_i mit $\mu_K = 10'000$ [CHF] , $\sigma_K = 2'000$ [CHF]
- Budget: Normalverteilte Zufallsvariable B, $\mu_B = 610'000$ [CHF], $\sigma_B = 50'000$ [CHF]

i. Grenzzustandsfunktion für die Überschreitung des Budgets:

$$M = B(normalverteilt) - K_{ges}(normalverteilt)$$

$K_{ges} = \sum_{i=1}^{50} K_i$ ist näherungsweise normalverteilt gemäss zentralem Grenzwertsatz.

ii. Zuverlässigkeitsindex & Wahrscheinlichkeit einer Budgetüberschreitung:

$$\mu_M = 610'000 - 50 \cdot 10'000 = 110'000$$
$$\sigma_M = \sqrt{(50'000)^2 + (2000)^2 \cdot 50} = \sqrt{2'700'000'000} = 51'961.5$$

$$\Rightarrow \beta = \frac{\mu_M}{\sigma_M} = \frac{110'000}{51'961.5} = \underline{\underline{2.117}}$$

$$\Rightarrow p_f = \Phi(-\beta) = \Phi(-2.117) = 1 - \Phi(2.12) = 1 - 0.98299698 = \underline{\underline{0.017}}$$

Aufgabe 4: Entscheidungsanalyse (25 Punkte)

a) *(6 Punkte)* **A-Priori-Entscheidungsanalyse**

A_1, erneuern: $(0.02\cdot 50+0.98\cdot 0 \text{ Mio.CHF})+5=6 \text{ Mio.CHF}$

A_2, belassen: $\left[0.25\cdot(0.3\cdot 50+0.7\cdot 0) \text{ Mio.CHF}\right]+\left[0.75\cdot(0.02\cdot 50+0.98\cdot 0) \text{ Mio.CHF}\right]$
$=4.5 \text{ Mio.CHF}$

➔ Aufgrund der A-Priori-Information würde sich die Ingenieurin für das Belassen der Lawinenverbauungen entscheiden.

b) *(3 Punkte)* **A-Posteriori-Analyse**

$$P(S|I)=\frac{P(I|S)\cdot P(S)}{P(I|S)\cdot P(S)+P(I|\overline{S})\cdot P(\overline{S})}=\frac{0.6\cdot 0.25}{0.6\cdot 0.25+0.3\cdot 0.75}=\frac{0.15}{0.375}=\underline{\underline{0.4}}$$

c) *(16 Punkte) :* **Prä-Posteriori-Entscheidungsanalyse**

<u>A priori Wahrscheinlichkeiten der Indikationen:</u>

$$P'[I]=P[I\,|\,S]\cdot P[S]+P\left[I\,|\,\overline{S}\right]\cdot P\left[\overline{S}\right]=0.6\cdot 0.25+0.3\cdot 0.75=0.375$$
$$P'\left[\overline{I}\right]=P\left[\overline{I}\,|\,S\right]\cdot P[S]+P\left[\overline{I}\,|\,\overline{S}\right]\cdot P\left[\overline{S}\right]=0.4\cdot 0.25+0.7\cdot 0.75=0.625$$

<u>A posteriori Wahrscheinlichkeiten der Zustände für jede Indikation:</u>

$$P''(S|I)=\frac{P(I|S)\cdot P(S)}{P'(I)}=\frac{0.6\cdot 0.25}{0.375}=\frac{0.15}{0.375}=0.4$$
$$P''(\overline{S}|I)=1-0.4=0.6$$

$$P''(S|\overline{I})=\frac{P(\overline{I}|S)\cdot P(S)}{P'(\overline{I})}=\frac{0.4\cdot 0.25}{0.625}=\frac{0.1}{0.625}=0.16$$
$$P''(\overline{S}|\overline{I})=1-0.16=0.84$$

Erwartete Kosten bei unterschiedlichen Indikationen

$$E\left[u \mid I\right] = \min\left\{6 \text{ Mio.CHF }; 0.4 \cdot 0.3 \cdot 50 + 0.6 \cdot 0.02 \cdot 50 \text{ Mio.CHF}\right\}$$
$$= \min\left\{6 \text{ Mio.CHF }; 6.6 \text{ Mio.CHF}\right\} = 6 \text{ Mio.CHF}$$

$$E\left[u \mid \overline{I}\right] = \min\left\{6 \text{ Mio.CHF }; 0.16 \cdot 0.3 \cdot 50 + 0.84 \cdot 0.02 \cdot 50 \text{ Mio.CHF}\right\}$$
$$= \min\left\{6 \text{ Mio.CHF }; 3.24 \text{ Mio.CHF}\right\} = 3.24 \text{ Mio.CHF}$$

Erwartete Kosten bei der Durchführung der visuellen Inspektion

$$E\left[u\right] = P\left[I\right] \cdot E\left[u \middle| I\right] + P\left[\overline{I}\right] \cdot E\left[u \middle| \overline{I}\right] + C$$
$$= 0.375 \cdot 6 + 0.625 \cdot 3.24 \text{ Mio.CHF} + C$$
$$=4.275 \text{ Mio.CHF} + C$$

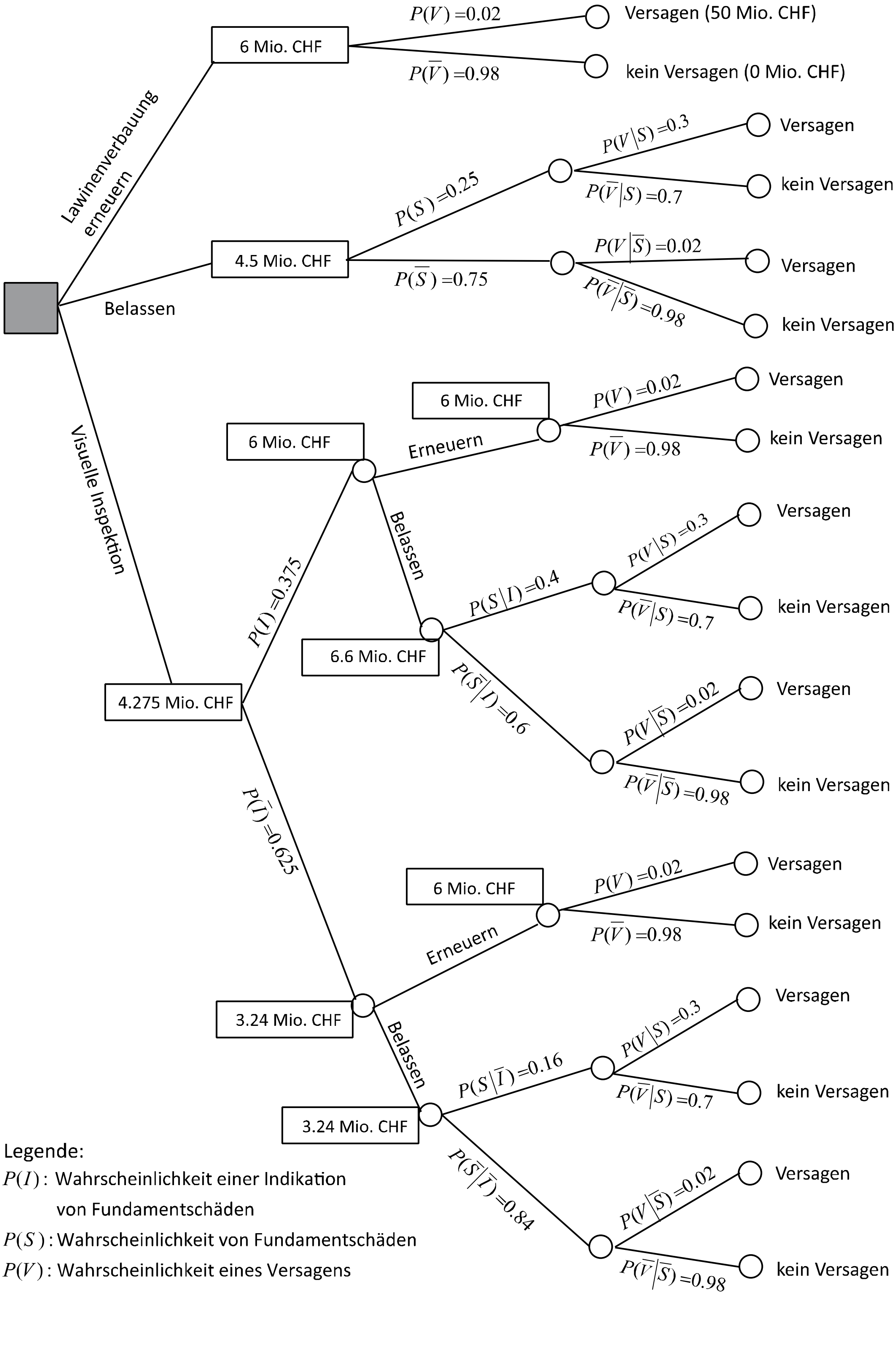
Prä-Posteriori-Entscheidungsanalyse

$$E\left[gesamt\right] = \min\left\{6 \text{ Mio.CHF}; 4.5 \text{ Mio.CHF}; 4.275 \text{ Mio.CHF} + C\right\}$$

$$C_{\max} = 4.5 - 4.275 = 0.225 \text{ Mio.CHF}$$

Rückschluss

Die visuelle Inspektion darf maximal 0.225 Mio.CHF kosten, damit sie sich lohnt.



Legende:

$P(I)$: Wahrscheinlichkeit einer Indikation von Fundamentschäden

$P(S)$: Wahrscheinlichkeit von Fundamentschäden

$P(V)$: Wahrscheinlichkeit eines Versagens

Aufgabe 5: Extremwerte, Zufallsprozesse (25 Punkte)

Maximale Tagestemperatur $T\left[{}^\circ C\right]$: Gumbel Max Verteilung ($u = 20$ und $\alpha = 0.5$)
Temperaturen unterschiedlicher Tage voneinander unabhängig

a) *(8 Punkte)* **Extremwertverteilung**

$$\begin{aligned} P(A_1) &= P(T > 33) = 1 - P(T \le 33) = 1 - F_X(33) \\ &= 1 - \exp(-\exp(-0.5(33-20))) = \underline{\underline{0.00150}} \end{aligned}$$

$$\begin{aligned} P(A_2) &= P(W > 33) = 1 - P(W \le 33) = 1 - (F_T(33))^7 \\ &= 1 - (\exp(-\exp(-0.5(33-20))))^7 = \underline{\underline{0.01047}} \end{aligned}$$

b) (*9 Punkte)* **Bernoulli-Versuche**

Mit $P(T > 33^\circ C) = 0.001$:

$$\begin{aligned} P(B_1) &= 1 - [P(2\,Tage) + P(1\,Tag) + P(0\,Tage)] \\ &= 1 - \left[\binom{31}{2} \cdot 0.001^2 \cdot (1-0.001)^{29} + \binom{31}{1} \cdot 0.001 \cdot (1-0.001)^{30} + (1-0.001)^{31} \right] \\ &= 1 - \left[0.00045 + 0.03008 + 0.96946\right] = \underline{\underline{0.00001}} \end{aligned}$$

$$P(B_2) = P(\text{erste drei Tage T>33°C}) = \left(P(T > 33^\circ C)\right)^3 = 0.001^3 = \underline{\underline{1 \cdot 10^{-9}}}$$

c) *(8 Punkte)* **Inhomogener Poissonprozess**

Poissonparameter:

$$u_1 = \int\limits_{Juli,Aug} \nu(t)dt = 15 \cdot 2 = 30$$

$$u_2 = \int\limits_{übrige} \nu(t)dt = \int\limits_{April} \nu(t)dt + \int\limits_{Mai,Jun} \nu(t)dt + \int\limits_{Sept} \nu(t)dt = 5 + 2 \cdot 10 + 5 = 30$$

Einsetzen in die Poissonverteilung:

$$P(C_1) = P_0(Juli, Aug) = \exp(-u_1) = \exp(-30) = \underline{\underline{9.36 \cdot 10^{-14}}}$$

$$P(C_2) = 1 - P_0(übrige) = 1 - \exp(-u_2) 1 - \exp(-30) = \underline{\underline{1 - 9.36 \cdot 10^{-14} \approx 1}}$$