

Musterlösung 12

1. a) Wir nehmen an, dass die Präferenzen der einzelnen Tennisspieler unabhängig voneinander und gleichverteilt sind, also ist X binomialverteilt mit $n = 20$ und p . Die Nullhypothese und Alternative sind

$$H_0 : p = 0.5 \quad H_A : p > 0.5.$$

Wir suchen also einen Verwerfungsbereich der die Nullhypothese verwirft, falls viel mehr als die Hälfte der Tennisspieler Darm-Saiten bevorzugen, also falls X gross ist. Der einzige Verwerfungsbereich der dies erfüllt ist der erste, also $\{15, 16, 17, 18, 19, 20\}$.

- b) Die Wahrscheinlichkeit eines Fehlers erster Art, also die Nullhypothese abzulehnen, obwohl sie stimmt ist

$$P_{H_0}[X \geq 15] = \sum_{k=15}^{20} \binom{20}{k} 0.5^k 0.5^{(20-k)} \approx 0.021.$$

Also liefert der in a) gewählte Verwerfungsbereich einen Test zum Niveau $\alpha = 0.05 \geq 0.021$. Wir erhalten

$$P_{H_0}[X \geq 14] \approx 0.058 > 0.05,$$

also ist der gewählte Verwerfungsbereich der grösstmögliche zum Niveau α .

- c) Ein Fehler zweiter Art passiert, wenn wir die Nullhypothese nicht verwerfen obwohl sie falsch ist, also

$$P_{p=p_1=0.6}[X \leq 14] = 0.874 \quad \text{und} \quad P_{p=p_2=0.8}[X \leq 14] = 0.196.$$

- d) Der P-Wert ist

$$P_{H_0}[X \geq 13] \approx 0.132,$$

die Nullhypothese wird also auf dem Signifikanzniveau 0.10 nicht verworfen.

2. Seien X_1, \dots, X_n iid Bernoulli(p)-verteilten Zufallsvariablen mit der Interpretation, dass $X_i = 1$, falls der i -ten Wurf Kopf ist und 0 sonst wobei $i = 1, \dots, n = 100$. Sei $X := \sum_{i=1}^n X_i$ die Anzahl Köpfe in den n Würfeln. Nach Voraussetzung ist X binomial verteilt mit Parametern n und p .

- a) zweiseitiger Test:

$$H_0 : p = p_0 = 1/2,$$

$$H_1 : p \neq p_0.$$

Unter H_0 ist $E[X] = np_0 = 50$ und $\sigma_X = \sqrt{np_0(1-p_0)} = 5$. Bestimmen des Verwerfungsbereiches: Für c_1, c_2 soll

$$\begin{aligned} 0.01 &\geq P_{H_0}[X \notin [c_1, c_2]] = 1 - P_{H_0}[X \in [c_1, c_2]] \\ &= 1 - P_{H_0} \left[\frac{c_1 - 50}{5} \leq \frac{X - E_{H_0}[X]}{\sigma_X} \leq \frac{c_2 - 50}{5} \right] \\ &\approx 1 - \Phi \left(\frac{c_2 - 50}{5} \right) + \Phi \left(\frac{c_1 - 50}{5} \right). \end{aligned}$$

Um den Verwerfungsbereich gross zu machen, muss das Intervall $[c_1, c_2]$ möglichst kurz werden. Wähle daher

$$\frac{c_1 - 50}{5} = -\frac{c_2 - 50}{5}.$$

Dann

$$0.01 \geq 1 - \Phi\left(\frac{c_2 - 50}{5}\right) + 1 - \Phi\left(\frac{c_2 - 50}{5}\right)$$

also $\Phi\left(\frac{c_2 - 50}{5}\right) \geq 0.995$ und $c_2 \geq 62.9$. Der Verwerfungsbereich ist also durch

$$K_{1\%} = [0, 37] \cup [63, 100]$$

gegeben und die Nullhypothese wird zum Signifikanzniveau 1% nicht verworfen.

b) Einseitiger Test:

$$H_0 : p = p_0 = 1/2,$$

$$H_1 : p \geq p_0.$$

Der Verwerfungsbereich durch die Ungleichung

$$0.01 \geq P_{H_0}[X > c] = 1 - \Phi\left(\frac{c - 50}{5}\right)$$

bestimmt. Also $c \geq 61.5$ und $K_{1\%} = [62, 100]$. Damit der Test zum 1% die Annahme einer fairen Münze nicht verwirft, darf höchstens 61 mal Kopf fallen.

c)

$$\frac{x}{n} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{x}{n}\left(1 - \frac{x}{n}\right) \frac{1}{n}} \approx 0.5 \quad \text{oder} \quad 0.7$$

also $[0.5, 0.7]$.

3. a) Da Herr User die theoretische Standardabweichung der Daten nicht kennt, wählt er einen t-Test. Seien X_1, \dots, X_6 die gemessenen Zeiten in ms mit arithmetischem Mittel \bar{X} .

Modellannahmen: X_1, \dots, X_6 sind unabhängig und normalverteilt mit unbekanntem Erwartungswert μ und unbekannter Standardabweichung σ .

Nullhypothese H_0 : $\mu = 5.8$.

Alternativhypothese H_A : $\mu < 5.8$.

Teststatistik T :

$$T = \frac{\bar{X} - 5.8}{s/\sqrt{6}}, \quad \text{wobei} \quad s = \sqrt{\frac{1}{6-1} \sum_{i=1}^6 (X_i - \bar{X})^2}.$$

Verteilung der Teststatistik unter Annahme von H_0 : Falls H_0 gilt, ist T t -verteilt mit 5 Freiheitsgraden.

Verwerfungsbereich: H_0 wird verworfen, falls $T \leq -3.365$ ist, da unter H_0 gemäss Tabelle gilt: $P[T \leq -3.365] \approx 0.01$.

Beobachtung: Wegen $\bar{X} = 5.60167$ und $s = 0.228772$ ist hier $T = -2.12354$, was nicht im Verwerfungsbereich liegt.

Testentscheid: H_0 wird nicht verworfen.

b) Es läuft alles wie bei Aufgabe nur, wobei diesmal jedoch $\bar{X} = 5.685$ und $s = 0.0605805$ sind. Daher ist hier $T = -4.64987$, was im Verwerfungsbereich liegt. Obwohl also der neue Mittelwert näher an 5.8 liegt als der alte, kann diesmal die Nullhypothese verworfen werden. Dies liegt daran, dass die geschätzte Standardabweichung s jetzt viel kleiner ist als vorher.

Bemerkung: Die Normalverteilungsannahme in Aufgabe a) scheint angesichts der Zahl 5.15 etwas fragwürdig zu sein.

4. a) Seien X_1, \dots, X_{10} die gemessenen Temperaturen mit arithmetischem Mittel \bar{X} . **Modellannahmen:** X_1, \dots, X_{10} sind unabhängig und normalverteilt mit unbekanntem Erwartungswert μ und Standardabweichung $\sigma = 0.5$.

Nullhypothese H_0 : $\mu = 20$.

Alternativhypothese H_A : $\mu \neq 20$.

Teststatistik T :

$$T = \frac{\bar{X} - 20}{0.5/\sqrt{10}}$$

Verteilung der Teststatistik unter Annahme von H_0 : Falls H_0 gilt, ist T standardnormalverteilt.

Verwerfungsbereich: H_0 wird verworfen, falls $|T| \geq 1.96$ ist, da unter H_0 gilt:
 $P[|T| \geq 1.96] \approx 0.05$.

Beobachtung: Wegen $\bar{X}_n = 20.543$ ist hier $T = 3.43$, was eindeutig im Verwerfungsbereich liegt.

Testentscheid: H_0 wird verworfen.

- b) Zu berechnen ist $P[|T| \geq 3.43] = 2P[T \geq 3.43] = 2(1 - \Phi(3.43)) = 0.0006$.

- c) **Modellannahmen:** Wie nehmen an, dass die X_i unabhängig und identisch verteilt sind.

Nullhypothese H_0 : Wir interpretieren die Aussage, dass die Klimaanlage richtig geeicht ist so, dass dann $p := P[X_i > 20] = 1/2$ gilt. Also lautet H_0 : $p = 1/2$.

Alternativhypothese H_A : $p \neq 1/2$.

Teststatistik T : Sei $Y_i = 1$, falls $X_i > 20$, und $Y_i = 0$, falls $X_i \leq 20$. Dann wählen wir als Teststatistik

$$T = Y_1 + \dots + Y_{10}.$$

Verteilung der Teststatistik unter Annahme von H_0 : Falls H_0 gilt, ist T binomialverteilt mit Parametern $n = 10$ und $p = 1/2$.

Beobachtung: Es wurde $T = 8$ beobachtet, da genau acht der zehn gemessenen Werte grösser als 20 sind.

P-Wert: Dies ist die Wahrscheinlichkeit unter der Annahme von H_0 , dass wir etwas beobachten, dass mindestens genauso stark im Sinne der Alternativhypothese von H_0 abweicht wie unsere tatsächliche Beobachtung, also

$$P_{H_0}[T \geq 8 \text{ oder } T \leq 2] = 2 \left(\binom{10}{8} + \binom{10}{9} + \binom{10}{10} \right) \left(\frac{1}{2} \right)^{10} = 0.109375.$$

Testentscheid: H_0 wird genau dann verworfen, wenn das Niveau α grösser ist als der P-Wert. Zum Niveau 5% z.B. wird H_0 anders als beim in a) behandelten Test nicht verworfen.

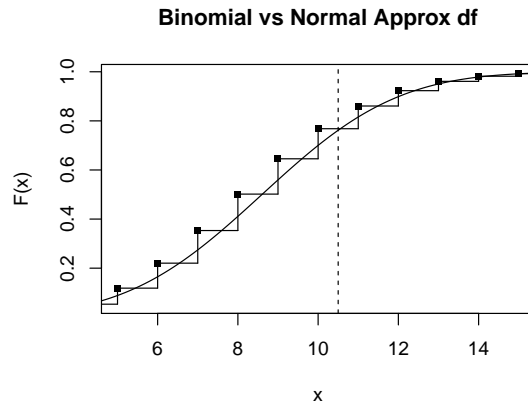
5. a) A =Fehler A tritt ein.
 B =Fehler B tritt ein.

$$\begin{aligned} P(\text{Fehler A oder Fehler B tritt ein}) &= P(A \cup B) \\ &= P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A) \cdot P(B) \\ &= 0.1 + 0.08 - 0.1 \cdot 0.08 = 0.172. \end{aligned}$$

- b) X =Anzahl unbrauchbarer Transistoren
 $X \sim \text{Bin}(50, 0.172)$.

- c) Wir verwenden die Normalapproximation für X , d.h. X ist näherungsweise $\mathcal{N}(np, np(1-p)) = \mathcal{N}(8.6, 7.12)$ -verteilt, woraus sofort folgt

$$P(X \leq 10) = P(X \leq 10.5) \approx \Phi\left(\frac{10.5 - 8.6}{\sqrt{7.12}}\right) = \Phi(0.712) = 0.76.$$



- d) X = Anzahl unbrauchbarer Transistoren

$$n = 100, \alpha = 0.01$$

$$H_0 : p_0 = 0.12$$

$$H_A : p > 0.12$$

$$\text{Unter } H_0: X \stackrel{\text{approx}}{\sim} \mathcal{N}(np_0, np_0(1-p_0)) = \mathcal{N}(12, 10.56)$$

i) Dann

$$\begin{aligned} P_{H_0}(X \geq 16) &= P_{H_0}(X > 15.5) = 1 - P_{H_0}(X \leq 15.5) \\ &\approx 1 - \Phi\left(\frac{15.5 - 12}{\sqrt{10.56}}\right) = 1 - \Phi(1.08) = 1 - 0.86 = 0.14 \end{aligned}$$

Da $0.14 > \alpha = 0.01$ wird H_0 beibehalten.

Alternativ:

ii) Verwerfungsbereich: c so, dass: $P_{H_0}(X \geq c) = \alpha = 0.01$,

$$P_{H_0}(X \geq c) = 1 - P_{H_0}(X < c) \approx 1 - \Phi\left(\frac{c - 12}{\sqrt{10.56}}\right) = 0.01,$$

d.h.

$$\frac{c - 12}{\sqrt{10.56}} = \Phi^{-1}(0.99) = 2.33 \quad \text{und somit} \quad c = 2.33\sqrt{10.56} + 12 = 19.57.$$

$$\Rightarrow K = \{20, \dots, 100\}.$$

Da $X \notin K$ wird H_0 nicht verworfen.

- e) $p_0 = 0.12$.

Wahrer Wert: $p_A = 0.16$.

Macht = $1 - P(\text{Fehler 2. Art}) = 1 - P_{p=p_A}(H_0 \text{ wird nicht verworfen})$.

Verwerfungsbereich in Abhängigkeit von n (schon in d) berechnet):

$$[c(n), \infty) \quad \text{wobei} \quad c(n) = 2.33\sqrt{np_0(1-p_0)} + np_0,$$

unter der Annahme $p = p_A$ ist X näherungsweise $\mathcal{N}(np_A, np_A(1-p_A))$ -verteilt. Gesucht ist n so, dass

$$\begin{aligned} 0.1 &\geq \mathbb{P}[X \notin K] = \mathbb{P}[X \leq c(n)] = \Phi\left(\frac{c(n) - np_A}{\sqrt{np_A(1-p_A)}}\right) \\ &= \Phi\left(\frac{2.33\sqrt{np_0(1-p_0)} + np_0 - np_A}{\sqrt{np_A(1-p_A)}}\right) \\ &= \Phi\left(\frac{2.33 \cdot 0.32 + \sqrt{n}(-0.04)}{0.37}\right), \end{aligned}$$

mit $\Phi^{-1}(0.1) = -\Phi^{-1}(0.9) = -1.28$ folgt daraus weiters

$$\begin{aligned} -1.28 &\geq \frac{0.76 + \sqrt{n}(-0.04)}{0.37} \\ -1.28 \cdot 0.37 - 0.76 &\geq -0.04\sqrt{n}, \end{aligned}$$

und schließlich $\sqrt{n} \geq 30.84$ bzw. $n \geq 951.1$.

6. a) Mit $\alpha = 0.05$, $\sigma^2 = 1$ und X_i den gemessenen Nägellängen ergibt sich

$$\frac{1}{n} \sum_{i=1}^n X_i \pm \sqrt{\frac{\sigma^2}{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

somit erhält man für die Länge des Vertrauensintervalls I:

$$|I| = \frac{2\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}{\sqrt{n}} = 0.5$$

Nach n aufgelöst: $n = \left(\frac{2 \cdot 1.96}{0.5}\right)^2 = 61.47$ Also muss man mindestens 62 Nägel messen.

b) Wegen dem zentralen Grenzwertsatz hat man:

$$\frac{\frac{1}{n} \sum_{i=1}^n X_i - p}{\sqrt{\frac{1}{n}p(1-p)}}$$

ist approximativ standardnormalverteilt. Mit der Normalapproximation erhalten wir also, mit $\alpha = 0.01$:

$$\frac{1}{n} \sum_{i=1}^n X_i \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{1}{n}p(1-p)}$$

Die Intervallgrenzen müssen von p unabhängig sein, deshalb maximieren wir $p(1-p)$. Das Maximum wird für $p = 0.5$ angenommen. Damit erhalten wir:

$$|I| = \frac{2\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}{2\sqrt{n}} = 0.01$$

Also $n = (100 \cdot 1.96)^2 = 38416$.

7. a) i) **t-Test**

- Modell: $(X_i)_{1 \leq i \leq 10} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\mu = E[X_i]$ unbekannt, $\sigma^2 = \text{Var}(X_i)$ ebenfalls unbekannt
- $H_0: \mu = \mu_0 = 20$
- $H_1: \mu \neq \mu_0$
- Teststatistik: $T := \frac{\bar{X} - \mu_0}{s_x} = \frac{\bar{X} - \mu_0}{s_x / \sqrt{n}}$
- Unter H_0 gilt: $T \sim t_9$. (Wir erwarten in diesem Fall also ein T in der Nähe von 0.)
- H_0 wird auf dem 5%-Niveau verworfen, falls $|T| > t_{9,0.975}$. Die kritischen Werte $t_{9,0.975}$ sind tabelliert. Für $n = 10$ haben wir $t_{9,0.975} = 2.26$.
- Mit unseren Daten beträgt die Teststatistik T

$$T = \frac{\bar{X} - \mu_0}{\frac{s_x}{\sqrt{n}}} = \frac{20.44 - 20}{\frac{0.813}{\sqrt{10}}} = 1.712 < t_{9,0.975} = 2.26$$

Das Gerät kann richtig geeicht sein. Die Nullhypothese wird nicht verworfen.

ii) **z-Test**

- a) Modell: $(X_i)_{1 \leq i \leq 10} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ unbekannt, σ^2 bekannt (im Gegensatz zu Fall i).
 b) $H_0: \mu = \mu_0 = 20$
 c) $H_1: \mu \neq \mu_0$
 d) Teststatistik: $Z := \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
 e) Unter H_0 gilt: $Z \sim \mathcal{N}(0, 1)$. (Wir erwarten in diesem Fall also ein Z in der Nähe von 0.)
 f) H_0 wird auf dem 5%-Niveau verworfen, falls $|Z| > 1.96$
 g) Mit unseren Daten beträgt die Teststatistik Z

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{20.44 - 20}{\frac{0.8}{\sqrt{10}}} = 1.739 < 1.96$$

Das Gerät kann richtig geeicht sein. Die Nullhypothese wird nicht verworfen.

- b) i) Vertrauensintervall beim t -Test: $[\bar{x} - t_{n-1}(\frac{\alpha}{2}) \frac{s_x}{\sqrt{n}}, \bar{x} + t_{n-1}(\frac{\alpha}{2}) \frac{s_x}{\sqrt{n}}]$
 Mit $\bar{x} = 20.44$, $t_9(0.025) = 2.262$, $s_x = 0.813$ und $n = 10$ ergibt dies: $[19.86, 21.02]$
 ii) Vertrauensintervall beim z -Test: $[\bar{x} - z(\frac{\alpha}{2}) \frac{\sigma_0}{\sqrt{n}}, \bar{x} + z(\frac{\alpha}{2}) \frac{\sigma_0}{\sqrt{n}}]$
 Mit $\bar{x} = 20.44$, $z(0.025) = 1.96$, $\sigma_0 = 0.800$ und $n = 10$ ergibt dies: $[19.94, 20.94]$

8. Quelle: J. Rice 1995, "Mathematical Statistics and Data Analysis", S. 313

- a) Als erste Vermutung erscheint die Poissonverteilung vernünftig, da 1. für jede Bakterienkolonie die Wahrscheinlichkeit, sich in einem bestimmten Quadrat zu befinden gleich klein sein dürfte, 2. die erwartete Anzahl Bakterienkolonien in jedem Quadrat gleich sein sollte, und 3. sich die Bakterienkolonien kaum gegenseitig beeinflussen.
 b) Sei X die Anzahl Kolonien pro Quadrätchen. Unter der Annahme, dass X Poisson-verteilt ist, gilt: $\lambda = \mathbb{E}[X]$. Als Schätzung verwenden wir:

$$\hat{\lambda} = \frac{\#\text{Kolonien}}{\#\text{Quadr.}} = \frac{\sum_{k=1}^{19} k \cdot N_k}{400} = 2.44$$

Die erwartete Anzahl Quadrätchen mit k Kolonien ist dann:

$$E_k = 400 \cdot \mathbb{P}[X = k] \approx 400 \cdot e^{-\hat{\lambda}} \frac{\hat{\lambda}^k}{k!} = \hat{E}_k$$

- c) Aus b) erhalten wir folgenden Abweichungen:

k	\hat{E}_k	$\hat{E}_k - N_k$
0	34.86	-21.14
1	85.07	-18.93
2	103.78	23.78
3	84.41	22.41
4	51.49	9.49
5	25.13	-1.87
6	10.22	1.22
7	3.56	-5.44
8	1.09	-3.91
9	0.29	-2.71
10	0.07	-1.93
≥ 11	0.02	-0.98

Wie man leicht sehen kann, sind die Abweichungen ziemlich gross. Deswegen scheint die Poissonverteilung in diesem Fall nicht zu passen. Um sicher zu gehen wird nun ein statistischer Test durchgeführt:

- d) Nullhypothese: Daten sind Poisson-verteilt mit Parameter $\lambda = 2.44$.
 Alternativhypothese: Daten sind *nicht* Poisson-verteilt mit Parameter $\lambda = 2.44$.
- e) Für die letzte Klasse hat man 11 Beobachtungen, wobei die Anzahl erwartete Beobachtungen nur

$$400 \cdot \mathbb{P}[X \geq 8] = 400(1 - \mathbb{P}[X \leq 7]) = 400(1 - \sum_{k=0}^7 \mathbb{P}[X = k]) \approx 1.47$$

beträgt.

Somit erhält man folgende Werte

k	$(\text{beob.} - \text{erwartet})^2 / \text{erwartet}$
0	12.81
1	4.21
2	5.45
3	5.95
4	1.75
5	0.14
6	0.15
7	8.30
≥ 8	62.64

Die Chiquadrat-Statistik beträgt dann

$$T = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}} \approx 100.4.$$

- f) Falls T zu gross ist, wird die Nullhypothese verworfen. Der Verwerfungsbereich ist also von der Form $V = [c_\alpha, \infty)$. Mit den c_α aus Tabelle in der Aufgabenstellung wird die Nullhypothese "Poissonverteilung" mit Parameter $\lambda = 2.44$ auf allen drei Niveaus verworfen.

9. Sei A_i der Gewichtsverlust im i -ten Durchgang bei Gewebetyp A und B_i derjenige von Gewebetyp B .

t -Test:

Modellannahmen: $A_i \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu_A, \sigma^2)$, $B_i \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu_B, \sigma^2)$, $i = 1, \dots, 10$, alle unabhängig.

Nullhypothese H_0 : $\mu_A = \mu_B$.

Alternativhypothese H_A : $\mu_A \neq \mu_B$.

Teststatistik T :

$$T = \frac{(\bar{A} - \bar{B}) / \sqrt{1/n + 1/n}}{\sqrt{(\sum (A_i - \bar{A})^2 + \sum (B_i - \bar{B})^2) / (n + n - 2)}}$$

Verteilung von T unter Annahme von H_0 : Falls H_0 gilt, ist T t -verteilt mit 18 Freiheitsgraden.

Verwerfungsbereich: H_0 wird verworfen, falls $|T|$ grösser ist als $t_{18}(0.975) = 2.10$.

Beobachtung: Es wurde beobachtet

$$\begin{aligned} \bar{A} &= 58.9 & \bar{B} &= 68.5 \\ \sum (A_i - \bar{A})^2 &= 918.9 & \sum (B_i - \bar{B})^2 &= 1226.5 \\ T &= -1.97 \end{aligned}$$

Testentscheid: H_0 wird nicht verworfen, da T nicht im Verwerfungsbereich liegt.

Mann-Whitney-Wilcoxon-Test:

Modellannahmen: $A_i \sim F$, $B_i \sim G$, alle unabhängig.

Nullhypothese H_0 : $F = G$.

Alternativhypothese H_A : $F \neq G$.

Teststatistik: $W = \sum_i \sum_j \mathbb{1}_{\{A_i < B_j\}}$

Verwerfungsbereich: H_0 wird gemäss Tabelle verworfen, falls $|W - \frac{n^2}{2}|$ mindestens 23 ist.

Beobachtung: $W = 73$, $|W - 50| = 23$.

Testentscheid: H_0 wird verworfen.

Diskussion:

Das Niveau ist beim Wilcoxon-Test immer korrekt, beim t -Test meist genähert richtig. Daher empfiehlt es sich, den Mann-Whitney-Wilcoxon-Test zu bevorzugen. Folglich sollte man die Hypothese, dass die beiden Gewebe bezüglich Abnutzung gleich sind, eher verwerfen.