

### Musterlösung zur Serie 3

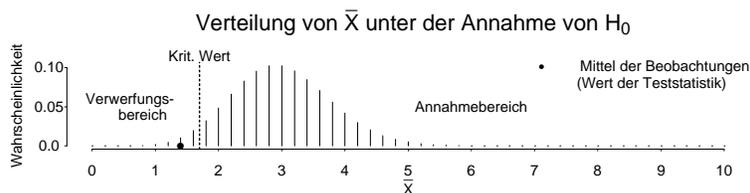
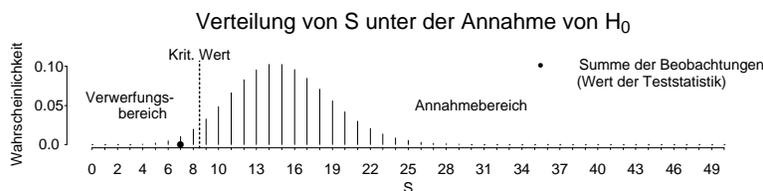
1. Die  $X_i$   $i = 1, \dots, n = 5$  bezeichnen die Anzahl Fasern in der  $i$ -ten Probe. Gemäss Annahme sind die  $X_i$  poissonverteilt mit je demselben Parameter  $\lambda_i = \lambda$  und unabhängig.

a) i) Nullhypothese  $H_0 : X_i \sim Pois(\lambda_0 = 3)$ , unabhängig. (=Produktionsprozess unverändert)

ii) Alternative  $H_A : \lambda < 3$ . (=Produktionsprozess verbessert)

Der Test ist einseitig durchzuführen, da wir nachweisen wollen, dass die Anzahl Fasern abgenommen hat (es interessiert nur eine Seite).

b) Gemäss dem Hinweis verwenden wir die Gesamtzahl der Fasern in 5 Proben als Zufallsvariable:  $S = \sum_{i=1}^5 X_i$ ; man kann zeigen, dass dann  $S \sim Pois(\tilde{\lambda} = n\lambda)$  (siehe Hinweis). Die Realisierung dieser Zufallsvariable ist  $s = x_1 + x_2 + x_3 + x_4 + x_5 = 1 + 0 + 2 + 1 + 3 = 7$ . Falls  $H_0$  richtig ist, ist  $\lambda = 3$  und somit  $\tilde{\lambda} = 15$ . Die Verteilung von  $S$  unter der Nullhypothese sieht also wie folgt aus:



c) Die kritische Grenze  $c$  erfüllt folgende Bedingung (dabei soll  $c$  möglichst gross sein):

$$P_0[S \leq c] = \sum_{k=0}^c \frac{15^k}{k!} \cdot e^{-15} \leq 0.05$$

Es ergeben sich folgende Wahrscheinlichkeiten:

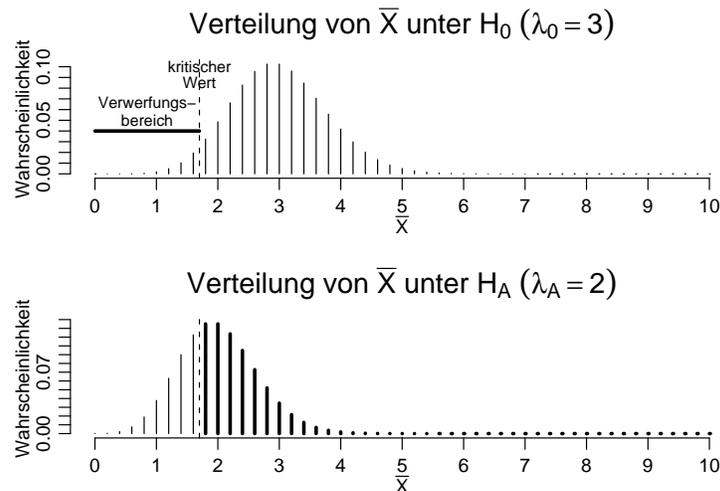
$s$	$P_0[S = s]$	$P_0[S \leq s]$	$s$	$P_0[S = s]$	$P_0[S \leq s]$
0	3.06e-07	3.06e-07	5	0.00194	0.00279
1	4.59e-06	4.89e-06	6	0.00484	0.00763
2	3.44e-05	3.93e-05	7	0.0104	0.018
3	0.000172	0.000211	8	0.0194	0.0374
4	0.000645	0.000857	9	0.0324	0.0699

Der Verwerfungsbereich  $K$  für ein Signifikanzniveau von 5% ist gegeben durch  $K = \{S \leq 8\}$  (vgl. Skizze unter Teilaufgabe b)).

- d) Aus den vorliegenden Daten berechnet sich  $s = 7$ . Dieser Wert liegt im Verwerfungsbereich der Nullhypothese. Die Nullhypothese kann demzufolge verworfen werden, es besteht ein signifikanter Unterschied zwischen dem alten und dem neuen Herstellungsverfahren.
- e) Unter  $H_A : \lambda_A = 2$  gilt  $S \sim Pois(n\lambda_A = 10)$ . Die Wahrscheinlichkeit  $\beta$  eines Fehlers 2. Art beträgt

$$\begin{aligned} P[H_0 : \lambda_0 = 3 \text{ beibehalten, obwohl } H_A : \lambda_A = 2 \text{ gilt}] &= P_{\lambda_A} [S > 8] \\ &= 1 - P_{\lambda_A} [S \leq 8] \\ &= 1 - \sum_{k=0}^8 \frac{10^k}{k!} \cdot e^{-10} \\ &= 1 - 0.3328 = 0.6672 \end{aligned}$$

Anschaulich entspricht  $\beta$  der Summe der fetten Wahrscheinlichkeitsstäbe im unteren Bild der folgenden Abbildung:



**Bemerkung:** Die Wahrscheinlichkeit eines Fehlers 2. Art ist in diesem Beispiel ziemlich hoch, d. h. wir können eine Verbesserung von 3 auf 2 kanzerogene Fasern pro Quadratmillimeter nur schwer nachweisen. Um die Situation zu verbessern, müsste der Hersteller den Stichprobenumfang erhöhen.

2. a) Das approximative Vertrauensintervall zum Niveau 0.05 ist (vgl. Skript, Abschnitt 3.3.3):

$$I = I(x) \approx x \pm 1.96\sqrt{x}$$

Das heisst  $\tilde{\lambda} \in 7 \pm 1.96\sqrt{7} = [1.81, 12.19]$ . Für  $\lambda = \tilde{\lambda}/5$  ist das zweiseitige Vertrauensintervall demnach  $[0.36, 2.44]$ .

- b) Das Vertrauensintervall von  $\lambda$  hat die gleiche Form wie das Vertrauensintervall von  $\tilde{\lambda}$  (um Faktor 5 gestreckt).

Um die genaue Form herauszufinden, müsste man für alle möglichen Werte von  $\tilde{\lambda}_0$  einen Hypothesentest  $H_0 : \tilde{\lambda} = \tilde{\lambda}_0$  gegen die *einseitige* Alternative  $H_A : \tilde{\lambda} < \tilde{\lambda}_0$  für die Beobachtung  $x = 7$  auf dem Signifikanzniveau  $\alpha$  durchführen. Alle Werte von  $\tilde{\lambda}_0$ , für die  $H_0$  nicht verworfen wird, liegen im *einseitigen*  $1 - \alpha$ -Vertrauensintervall von  $\tilde{\lambda}$ .

Wir haben schon gesehen, dass der obige Test für  $\tilde{\lambda}_0 = 15$  auf dem 5% Signifikanzniveau verworfen wird. Also liegt 15 nicht im 95%-Vertrauensintervall von  $\tilde{\lambda}$ . Für alle anderen Werte von  $\tilde{\lambda}$  müsste

man nun eigentlich eine ähnliche Rechnung durchführen. Das ist natürlich (ohne Computer) zu aufwändig und wir werden uns deshalb mit einer qualitativen Überlegung begnügen:

Für unseren Test ist der Annahmehereich immer von der Form  $K = \{a + 1, a + 2, \dots\}$ ; dabei ist  $a$  so gewählt, dass  $P_{H_0}[S \leq a] = \alpha$ . Wenn  $\tilde{\lambda}$  nun immer grösser wird, dann wird die ganze Wahrscheinlichkeitsverteilung von  $S$  nach rechts gestreckt, und daher wird auch  $a$  immer grösser. Irgendwann wird  $a$  grösser als  $x$  sein und dann wird  $H_0$  verworfen, d.h., das entsprechende  $\tilde{\lambda}$  gehört nicht mehr ins Vertrauensintervall. Wir haben also gesehen, dass sehr grosse Werte von  $\tilde{\lambda}$  nicht ins Vertrauensintervall kommen.

Wenn umgekehrt  $\tilde{\lambda}$  immer kleiner wird, so wird die Wahrscheinlichkeitsverteilung nach links gestaucht (wird aber nie negativ) und auch  $a$  wird immer kleiner. Somit ist für hinreichend kleine Werte von  $\tilde{\lambda}$  sicher  $x > a$ . D.h.,  $H_0$  wird nicht abgelehnt und  $\tilde{\lambda}$  kommt ins Vertrauensintervall. Wir haben also gesehen, dass sehr kleine Werte von  $\tilde{\lambda}$  ins Vertrauensintervall kommen.

Das Vertrauensintervall ( $\alpha$  beliebig) für  $\tilde{\lambda}$  muss also von der Form  $[0, c]$  sein. Zudem wissen wir, dass obiger Test für  $\alpha = 0.05$  und  $\tilde{\lambda} = 15$  verworfen wird. D.h., in dem 95% Vertrauensintervall ist  $c < 15$ . Der Computer liefert für das 95% Vertrauensintervall  $[0, 13.15]$ .