Solutions to Exercise Sheet 3

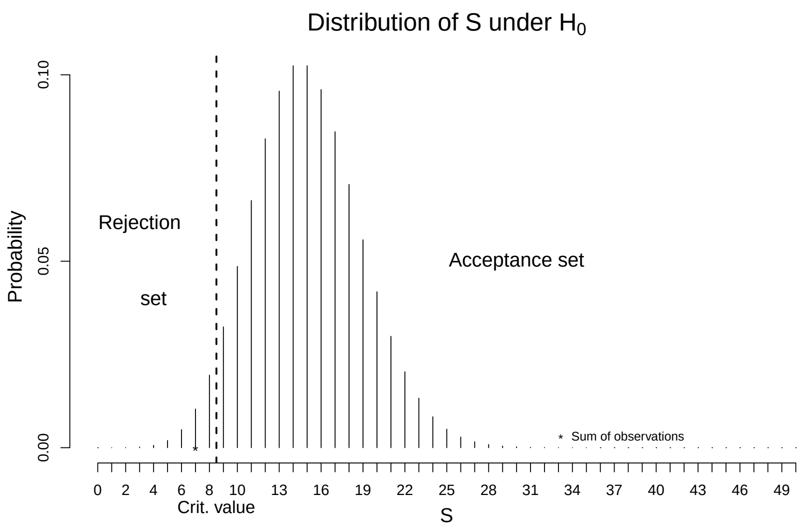
1. The numbers X_i $i = 1, \ldots, n = 5$ denote the amount of fibres in the i-th sample. By the assumption made, the X_i all independently follow a Poisson distribution with the same parameter $\lambda_i = \lambda$.

a) i) Null hypothesis: $H_0 : X_i \sim Pois(\lambda_0 = 3)$, independent. (=The outcome of the production method stays the same)

ii) Alternative hypothesis: $H_A : \lambda < 3$. (=The outcome of the production method is an improvement)

We should carry out a one-sided test, as we are looking for a decrease in the number of fibres (only one direction is of interest).

b) By the hint, we can consider the number of fibres in all 5 samples together to be a random variable: $S = \sum_{i=1}^{5} X_i$. It can be shown that $S \sim Pois(\tilde{\lambda} = n\lambda)$ (see the hint). For this random variables, wehave the realisation $s = x_1 + x_2 + x_3 + x_4 + x_5 = 1 + 0 + 2 + 1 + 3 = 7$. If H_0 is true, $\lambda = 3$ and thereby $\tilde{\lambda} = 15$. Under the null hypothesis, then, the distribution of S looks as follows:



c) The critical value c satisfies the following condition (for which c should be as large as possible):

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

The probabilities that follow are:

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

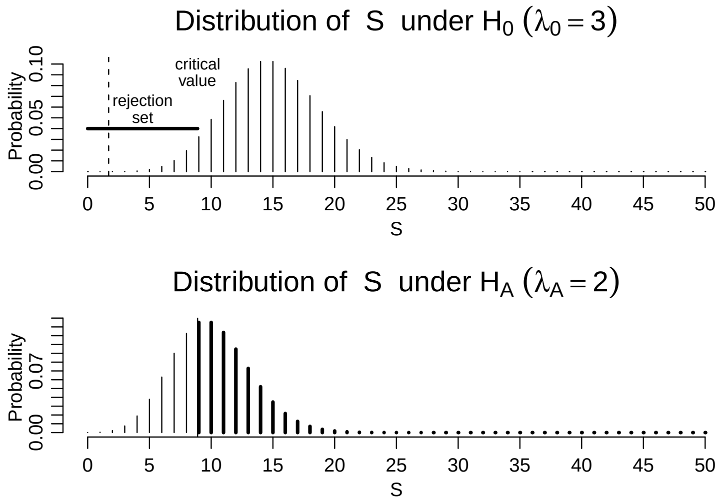
The rejection set K at level 5% is given by
$K = \{S \leq 8\}$ (cf. the sketch in Part b)).

d) The data yield $s = 7$. This value lies in the rejection set for the null hypothesis. Therefore, we can reject the null hypothesis, and a significant difference between the old and the new production methods does exist.

e) Under the alternative hypothesis $H_A : \lambda_A = 2$, we have $S \sim Pois(n\lambda_A = 10)$. The probability of a Type II error, β, is

$$\begin{aligned}
\mathrm{P}\left[H_0 : \lambda_0 = 3 \text{ is kept, even though } H_A : \lambda_A = 2 \text{ is true}\right] &= \mathrm{P}_{\lambda_A}\left[S > 8\right] \\
&= 1 - \mathrm{P}_{\lambda_A}\left[S \leq 8\right] \\
&= 1 - \sum_{k=0}^{8} \frac{10^k}{k!} \cdot e^{-10} \\
&= 1 - 0.3328 = 0.6672
\end{aligned}$$

β corresponds to the sum of the wide probability bars in the lower half of the following figure:



Comment: The probability of a Type II error is quite high in this example, i.e. it is quite difficult to show that the number of carcinogenic fibres has gone from 3 to 2. To improve the situation, the manufacturer would have to raise the sample size.

2. a) The approximate confidence interval at level 0.05 is (cf. Section 3.3.3 of the lecture notes):

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

Thus $\tilde{\lambda} \in 7 \pm 1.96\sqrt{7} = [1.81, 12.19]$. The two-sided confidence interval for $\lambda = \tilde{\lambda}/5$ is thus $[0.36, 2.44]$.

b) The confidence interval for λ has exactly the same form as the confidence interval for $\tilde{\lambda}$ (albeit stretched by a factor of 5).
To find out the precise form, we would need to perform a test of the hypothesis $H_0 : \tilde{\lambda} = \tilde{\lambda}_0$ against the *one-sided* alternative $H_A : \tilde{\lambda} < \tilde{\lambda}_0$ given the observation $x = 7$ and the significance level α. All those values of $\tilde{\lambda}_0$ that do not lead to the rejection of H_0 lie in the *one-sided* $1 - \alpha$ confidence interval for $\tilde{\lambda}$.
We have already seen that the test given above leads to the rejection of the null hypothesis $\tilde{\lambda}_0 = 15$ at a significance level of 5%. Thus 15 does not lie in the 95% confidence interval for

$\tilde{\lambda}$. In principle, we would now need to carry out similar calculations for all other values of $\tilde{\lambda}$. Without a computer, this is of course far too tedious, and we shall thus content ourselves with some qualitative reasoning:

Our test always has an acceptance set of the form $K = \{a+1, a+2, \ldots\}$; here a is chosen so as to ensure $P_{H_0}[S \leq a] = \alpha$. If now $\tilde{\lambda}$ is increased, the whole distribution of S is stretched out to the right, and so a increases in turn. At some point, a will exceed x; at that point, H_0 will be rejected, i.e. the corresponding value of $\tilde{\lambda}$ no longer belongs to the confidence interval. Thus we have seen that sufficiently large values of $\tilde{\lambda}$ do not lie in the confidence interval.

Conversely, if $\tilde{\lambda}$ is made ever smaller, the distribution is squashed up to the left (but never gets negative) and a too decreases. Thus for sufficiently small values of $\tilde{\lambda}$, we can be certain that $x > a$. Then H_0 is longer rejected, and $\tilde{\lambda}$ enters the confidence interval. Thus we have seen that small values of $\tilde{\lambda}$ do lie in the confidence interval.

Thus the confidence interval for $\tilde{\lambda}$ (for an arbitrary α) must be of the form $[0, c]$. We furthermore know that the above test is rejected when $\alpha = 0.05$ and $\tilde{\lambda} = 15$. I.e. the upper end of the 95% confidence interval is $c < 15$. This 95% confidence interval is given by the computer as $[0, 13.15]$.