Probability and Statistics

Exercise sheet 12

Exercise 12.1 Suppose that $X_1, \dots, X_n$ form a random sample from a Poisson distribution for which the mean λ is unknown. Determine the maximum likelihood estimator for λ.

Solution 12.1 Given the parameter λ, the probability weight function of the Poisson distribution is

$$\mathbb{P}_\lambda[X = k|\lambda] = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Hence the MLE is the value λ which maximizes, for $x_i = X_i$,

$$L(\lambda; x_1, \dots, x_n) = \prod_{i=1}^n e^{-\lambda}\frac{\lambda^{x_i}}{x_i!} = \frac{(e^{-\lambda}\lambda^{\bar{x}_n})^n}{x_1! \dots x_n!},$$

where $\bar{x}_n = (x_1 + \cdots + x_n)/n$. We need to find the λ which maximizes

$$g(\lambda) = e^{-\lambda}\lambda^{\bar{x}_n} = \exp(-\lambda + \bar{x}_n \log \lambda),$$

so we compute

$$g'(\lambda) = (-1 + \bar{x}_n/\lambda)g(\lambda)$$

and set this to 0. The maximum of g is reached when $\lambda = \bar{x}_n$. Thus for the sample $X_1, \dots, X_n$, the MLE for λ is $\bar{X}_n$.

Exercise 12.2 In the year 1910, Rutherford observed the radioactive decay of a substance in $n = 2608$ time intervals, each of 7.5 seconds. We use almost the same notation as Example 1.6.8 in the lecture notes: $\tilde{n}_k$ is the number of intervals with exactly k decays. We want to match a distribution to these data, and our null hypothesis H_0 is that the number of decays per interval is Poisson-distributed with unknown parameter λ.

Rutherford's experiments resulted in the following table:

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	≥ 15
$\tilde{n}_k$	57	203	383	525	532	408	273	139	45	27	10	4	0	1	1	0

Table 1: Original table.

In order to fulfil the rule of thumb when a χ^2 asymptotic is an appropriate approximation, we merge the rare cases in the following way:

k	0	1	2	3	4	5	6	7	8	9	10	11	≥ 12
n_k	57	203	383	525	532	408	273	139	45	27	10	4	2

Table 2: Merged table.

(a) Do a χ^2 test with the given data. (You can use appropriate approximations.)

Hint: Remember what you have learned about χ^2 tests in the lecture. Use Exercise 12.1.

(b) Do a χ^2 test with the given data for the alternative null hypothesis H_0': The number of decays per interval is Poisson-distributed with (exogenously given) parameter $\lambda' = 3.87$.

(c) Do you think a Poisson distribution is a plausible model? Do you think H_0' is plausible?

Solution 12.2

(a) Our null hypothesis has a composite form parametrized by $\eta = \lambda$, which gives

$$\Theta_0 = \{\theta_0(\lambda) = (\theta_{0,0}(\lambda), \dots, \theta_{0,12}(\lambda)) : \lambda \in (0, \infty)\},$$

where

$$\theta_{0,i}(\lambda) = \begin{cases} e^{-\lambda}\frac{\lambda^i}{i!}, & 0 \le i \le 11, \\ \sum_{j=12}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}, & i = 12. \end{cases}$$

In section 7.2.5 of the lecture notes, the asymptotics for a χ^2 test are given when we use the realized value of the MLE given as in the lecture notes by

$$\hat{\lambda}(\omega) = \underset{\lambda > 0}{\arg\max} \sum_{i=0}^{12} n_i \log \theta_{0,i}(\lambda).$$

We solve this by setting the first derivative to zero:

$$\begin{aligned}
\frac{\partial}{\partial \lambda} \sum_{i=0}^{12} n_i \log \theta_{0,i}(\lambda) &= 0 \\
\frac{\partial}{\partial \lambda}\left(\sum_{i=0}^{11} n_i \log\left(e^{-\lambda}\frac{\lambda^i}{i!}\right) + n_{12}\log\left(\sum_{j=12}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}\right)\right) &= 0 \\
\frac{\partial}{\partial \lambda}\left(\sum_{i=0}^{11} n_i\left(-\lambda + i\log(\lambda) - \log(i!)\right) + n_{12}\left(-\lambda + \log\left(\sum_{j=12}^{\infty}\frac{\lambda^j}{j!}\right)\right)\right) &= 0 \\
\sum_{i=0}^{11} n_i\left(-1 + \frac{i}{\lambda}\right) + n_{12}\left(-1 + \frac{\sum_{j=12}^{\infty}\frac{j\lambda^{j-1}}{j!}}{\sum_{j=12}^{\infty}\frac{\lambda^j}{j!}}\right) &= 0 \\
\frac{1}{\lambda}\left(\sum_{i=0}^{11} n_i i + n_{12}\frac{\sum_{j=12}^{\infty}\frac{j\lambda^{j-1}}{j!}}{\sum_{j=12}^{\infty}\frac{\lambda^j}{j!}}\right) &= \sum_{i=0}^{12} n_i.
\end{aligned}$$

This leads to the fixed point equation[1]

$$\hat{\lambda}(\omega) = \frac{\sum_{i=0}^{11} i n_i + \frac{\sum_{i=12}^{\infty}\frac{i\hat{\lambda}(\omega)^i}{i!}}{\sum_{i=12}^{\infty}\frac{i}{i!}} n_{12}}{\sum_{i=0}^{12} n_i}.$$

This can be easily solved numerically by a fixed point iteration. A suitable starting point for the iterative algorithm is the realized value of the MLE based on the original data, becasue this can be easily computed with the help of Exercise 12.1 resulting in

$$\hat{\lambda}_0(\omega) = \bar{X}_n(\omega) = \frac{\sum_{i=0}^{14} i\tilde{n}_i}{\sum_{i=0}^{14}\tilde{n}_i} = \frac{10097}{2608} = 3.8715.$$

[1]This formula might be more interpretable if we rewrite it as

$$\hat{\lambda}(\omega) = \frac{\sum_{i=0}^{11} i n_i + \mathbb{E}_{\hat{\lambda}(\omega)}[X_1 | X_1 \ge 12] n_{12}}{\sum_{i=0}^{12} n_i} = \mathbb{E}_{\hat{\lambda}(\omega)}[\bar{X}_n | N_1 = n_1, \dots, N_{12} = n_{12}] = \mathbb{E}_{\hat{\lambda}(\omega)}[\bar{X}_n | N_1, \dots, N_{12}](\omega)$$

by using the notation from Exercise 12.1 and the lecture notes, but this is not necessary for this exercise.

After two iterations we get $\hat{\lambda}(\omega) = 3.8707$. (More iterations would not result in visible changes with the shown number of decimals.) We see that the two values $\hat{\lambda}_0, \hat{\lambda}$ based on Tables 1 and 2 are different, but in this example they are approximately the same, $\hat{\lambda}_0(\omega) \approx 3.87 \approx \hat{\lambda}(\omega)$. So we continue with the approximation $\hat{\lambda}(\omega) = 3.87$. The following computations would be a bit more tedious, but can be easily done with the help of R (see `https://www.kaggle.com/jakobheiss/sol2-2/edit`). Alternatively one can avoid these computations by using the third row of the table in Example 1.6.8 in the lecture notes (by approximating $np_i, \ldots, np_{12}$ by integers—this approximation is not done for mathematical reasons but to save your time typing the numbers into your calculator). We continue to calculate with this approximation in this text (check the R-code for more accurate computations; we will discuss in the exercises class why the results are quite different.). Now, we can compute the test statistic

$$T(\omega) = \sum_{i=0}^{12} \frac{\left(n_i - n\theta_{0,i}\left(\hat{\lambda}(\omega)\right)\right)^2}{n\theta_{0,i}\left(\hat{\lambda}(\omega)\right)} = 13.9.$$

Since we estimated $r = 1$ parameter, under the null hypothesis, T is asymptotically χ^2-distributed with $13 - r - 1 = 11$ degrees of freedom, i.e. $T \sim \chi^2_{11}$. So the approximate realized value of the P-value is

$$\pi(\mathbf{X}(\omega)) \approx 1 - \chi^2_{11}(13.9) = 0.24.$$

This is usually not seen as a reason to reject the null hypothesis.

(b) As we are not estimating any parameters here, the test statistic is $T' = \sum_{i=0}^{12} \frac{(N_i - n\theta_{0,i}(\lambda'))^2}{n\theta_{0,i}(\lambda')}$, with realized value $T'(\omega) = \sum_{i=0}^{12} \frac{(n_i - n\theta_{0,i}(\lambda'))^2}{n\theta_{0,i}(\lambda')} = 13.9$, and T' is approximately χ^2_{12}-distributed under H'_0. The approximate realized value of the P-value is therefore

$$\pi'(\mathbf{X}(\omega)) \approx 1 - \chi^2_{12}(13.9) = 0.307.$$

This is usually not seen as a reason to reject the null hypothesis.

(c) From a physicist's perspective, a Poisson distribution appears to be very reasonable, because of properties such as Proposition 1.6.10 in the lecture notes. The χ^2 test does not suggest to reject the null hypothesis, so all in all it is very plausible that the true distribution is Poisson. If you suggest another null hypothesis, a χ^2 test based on our data might reject it.

$\pi < \pi'$ cannot be interpreted that H'_0 is more plausible than H_0. (One could come up with this wrong impression if one incorrectly interpreted the P-value π as the probability that H_0 were true—see the discussion in the lecture notes.) Actually, it is the other way around, H_0 is more plausible than H'_0, since $\Theta'_0 \subset \Theta_0$. Indeed, intuitively (or from a Bayesian perspective with a reasonable prior), one could say that it is almost impossible that H'_0 is *exactly* true. But it is plausible that H'_0 is approximately true and we have no strong reason to reject it as reasonable approximation for the truth.

Exercise 12.3 Let X be a normal random variable with $\mathbb{E}[X] = m$ and $\text{Var}[X] = \sigma^2 = 0.0014^2$. Let also X_i for $i = 1, \ldots, n$ be i.i.d. random variables that share the same distribution with X. The following 12 realisations x_i of the random variables X_i were recorded:

1.00781 1.00646 1.00801 1.00833 1.00738 1.00687
1.00783 1.00936 1.00564 1.00543 1.00794 1.01060

(a) Perform a statistical test at a level of confidence $\alpha = 5\%$ for the null hypothesis $H_0 : \mu = 1.0085$ against the alternative hypothesis $H_A : \mu = 1.008$.

(b) Calculate the power of the test from part (a).

(c) What happens to the power calculated in part (b) when the alternative hypothesis is changed to $H'_A \;:\; \mu = 1.007$?

Solution 12.3

(a) The null and alternative hypotheses are

$$H_0 : \mu = \mu_0 = 1.0085 \quad \text{and} \quad H_A : \mu = \mu_A = 1.008.$$

We can use the test statistic $T = \bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$, which has under H_0 distribution $\mathcal{N}(\mu_0, \sigma/\sqrt{n})$. The rejection rule is that the null hypothesis is rejected if $T \leq c_\alpha$, with c_α determined by

$$\mathbb{P}_{\mu_0}[T \leq c_\alpha] = \mathbb{P}_{\mu_0}\left[Z \leq \sqrt{n}\frac{c_\alpha - \mu_0}{\sigma}\right] = \Phi\left(\sqrt{n}\frac{c_\alpha - \mu_0}{\sigma}\right) = \alpha,$$

where Z has a standard normal distribution under H_0. Hence, using the table, we obtain $\sqrt{n}\frac{c_\alpha - \mu_0}{\sigma} = -1.645$, thus $c_\alpha \approx 1.0078$. We find that the realized value of the test statistic is $T(\omega) = \bar{x}_n = \frac{\sum_{i=1}^n x_i}{n} \approx 1.0076$. Consequently, since $T(\omega) < c_\alpha$, we reject the null hypothesis H_0.

(b) The power of the test is equal to the probability that the null hypothesis is rejected given that the alternative hypothesis is true, i.e.,

$$\beta = \mathbb{P}_{\mu_A}[T \leq c_\alpha] = \mathbb{P}_{\mu_A}\left[Z \leq \sqrt{n}\frac{c_\alpha - \mu_A}{\sigma}\right] = \Phi\left(\sqrt{n}\frac{c_\alpha - \mu_A}{\sigma}\right),$$

where Z has a standard normal distribution under H_A. We have calculated in part (a) $c_\alpha \approx 1.0078$; thus we get

$$\beta = \mathbb{P}_{\mu_A}\left[Z \leq \sqrt{n}\frac{c_\alpha - \mu_A}{\sigma}\right] = \Phi\left(\sqrt{n}\frac{c_\alpha - \mu_A}{\sigma}\right) \approx \Phi(-0.4949) = 1 - \Phi(0.4949) = 0.3085$$

(c) Changing to $\mu'_A = 1.007$ leads then to a higher power of the test. A similar calculation provides

$$\beta = \mathbb{P}_{\mu'_A}\left[Z \leq \sqrt{n}\frac{c_\alpha - \mu'_A}{\sigma}\right] = \Phi\left(\sqrt{n}\frac{c_\alpha - \mu'_A}{\sigma}\right) \approx \Phi(1.9795) = 0.9761,$$

which is much better, as one expects.

Exercise 12.4 In a study on the reliability of ball-bearings (in German: Kugellager), two samples of 10 pieces each of two different types of ball-bearings were tested. The number of rotations (in millions) until break-down were

type I	3.03	5.53	5.60	9.30	9.92	12.51	12.95	15.21	16.04	16.84
type II	3.19	4.26	4.47	4.53	4.67	4.69	12.78	6.79	9.37	12.75

Before the realization of this study, it was not clear which type was more reliable.

(a) Are we dealing with a paired sample? Please explain your answer.

(b) Perform a t-test for the null hypothesis “the expected number of rotations until break-down is the same for the two types of ball-bearings” with level 5%. (What are the model assumptions of a t-test?)

(c) Which other test would be a better alternative (fewer model assumptions and usually better power)? You can run that test in R.

Hint: (Clicking the following link will reveal the solution of (c).) You can find the R-code at `https://www.kaggle.com/jakobheiss/probstat2020-ex12-4/edit`.

Solution 12.4

(a) This is not a paired sample—the only connection between the data is their numbering, which does not give natural pairs.

(b) The model is given by $X_1, \ldots, X_{10}$ i.i.d. $\sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_1, \ldots, Y_{10}$ i.i.d. $\sim \mathcal{N}(\mu_Y, \sigma^2)$, where μ_X, μ_Y and σ are unknown and the X_i, Y_j are all independent. The null and alternative hypothesis are given by

$$H_0 : \mu_X = \mu_Y \quad \text{und} \quad H_A : \mu_X \neq \mu_Y.$$

The test statistic is

$$T := \frac{\bar{X}_n - \bar{Y}_m}{S_{\text{pool}} \sqrt{1/n + 1/m}},$$

where the estimator S_{pool} for σ is given by

$$S_{\text{pool}} = \sqrt{\frac{1}{n+m-2} \left(\sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + \sum_{k=1}^{m} (Y_k - \bar{Y}_m)^2 \right)},$$

and under H_0, T is t-distributed with $n + m - 2 = 18$ degrees of freedom. With a level of 5%, the null hypothesis will be rejected when $|T| > t_{18,0.975} = 2.101$. From the data, we obtain $\bar{x}_{10} = 10.693, \bar{y}_{10} = 6.75$ and $S_{pool}(\omega) = 4.255$, so $T(\omega) = 2.0723$, i.e., H_0 is not rejected. (The realized value of the P-value is $\pi((\mathbf{X}, \mathbf{Y})(\omega)) = 2\,(1 - t_{18}\,(2.0723)) = 0.053$, so we were close to rejecting.)

(c) The 2-sample Wilcoxon test (also known as Mann–Whitney U-test) does not assume normal distributions. The null hypothesis of this test is that all samples $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ are i.i.d. with respect to an arbitrary (unknown) continuous distribution F. The following R-code

```
wilcox.test(c(3.03, 5.53, 5.60, 9.30, 9.92, 12.51, 12.95, 15.21, 16.04, 16.84),
            c(3.19, 4.26, 4.47, 4.53, 4.67, 4.69, 12.78, 6.79, 9.37, 12.75),
            paired=FALSE)
```

results in a realized P-value of 0.063.

If you have feedback regarding the exercise sheets, please send a mail to Jakob Heiss.