

PHYS 849: Assignment 5

Jake Bauer

May 24, 2017

Problem 1

(a)

We are asked to go through the procedure of testing a hypothesis about the mean of this distribution from the data. To explore this question, we generate 10 random data points from a Gaussian with $\mu = 100$ and $\sigma = 10$. The log-likelihood function for a general Gaussian PDF of unknown mean μ and known standard deviation σ is given by,

$$\ln \mathcal{L} = -\frac{n}{2\sigma^2} \mu [\mu - 2\bar{x}] \quad (1)$$

where \bar{x} is the data sample mean, and where we have dropped factors that do not depend on the unknown quantities. Such terms would just correspond to constant shifts of the likelihood, and are consequently negligible. We are then asked to test:

- H_0 : $\mu = 100$
- H_1 : $\mu \neq 100$

where H_0 and H_1 are the null and alternate hypotheses. To compare the relative likelihood of these hypotheses, we construct a test statistic,

$$\ln \Lambda = \ln \frac{\mathcal{L}_0}{\mathcal{L}_1}, \quad (2)$$

or the difference in log-likelihoods. For the null hypothesis log-likelihood, we take Eq. (1) with $\mu = 100$. For H_1 , we know that \bar{x} is an unbiased and consistent estimator of μ , so we take $\mu = \bar{x}$. If $\mu = \bar{x}$, the test statistic is minimized and the hypotheses are equivalent. Alternatively, a high $\ln \Lambda$ means the null hypothesis is preferred and a low one means the opposite. To test these hypotheses, we draw 10^4 samples of 10 from the null hypothesis distribution in addition to the original data set. These form the basis of a Monte Carlo analysis in which the test statistic is computed for a range of \bar{x} . This distribution plotted in Figure 1, and the value generated by the original sample is shown by the position of the dashed line. The p-value of this observation is 0.3618, so we fail to reject the null hypothesis.

(b)

We will follow the same procedure to test,

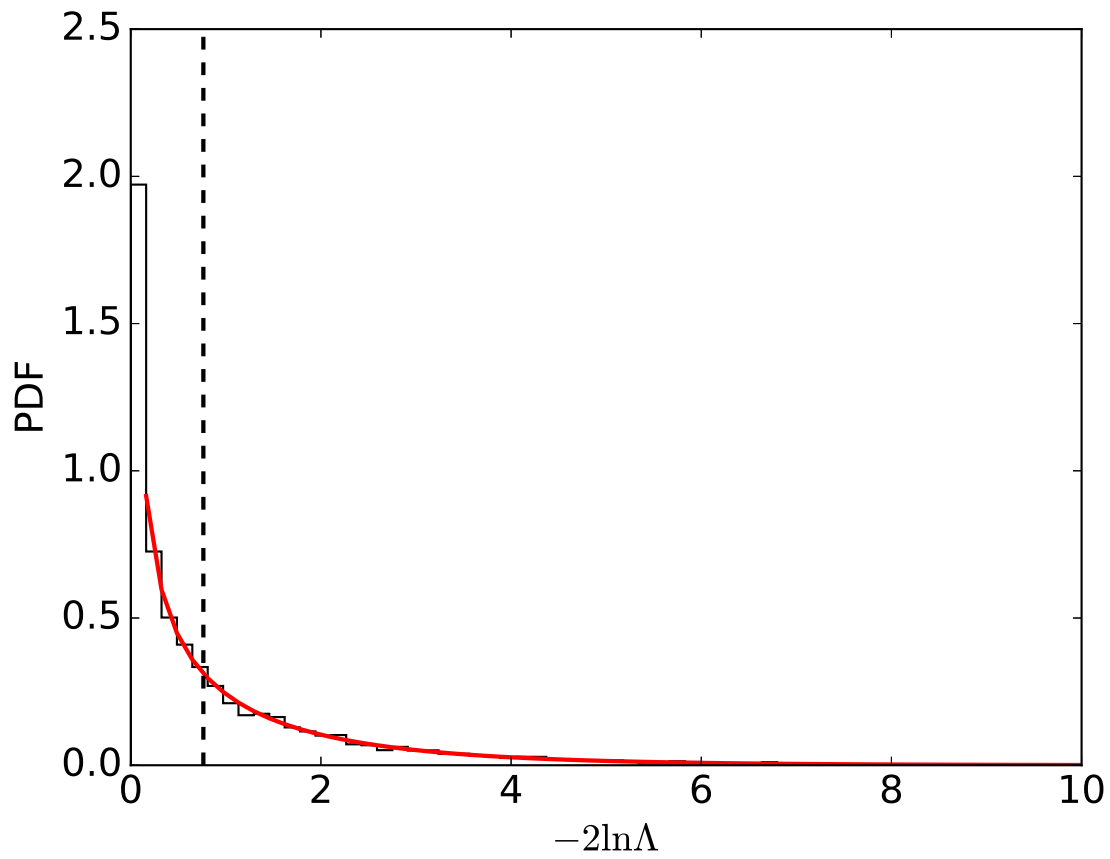


Figure 1: The test statistic for $H_0 = 100$, $H_1 \neq 100$. The prediction from Wilk's Theorem is overlaid. The sample data set's value of $-2\ln\Lambda$ is denoted by the dashed line.

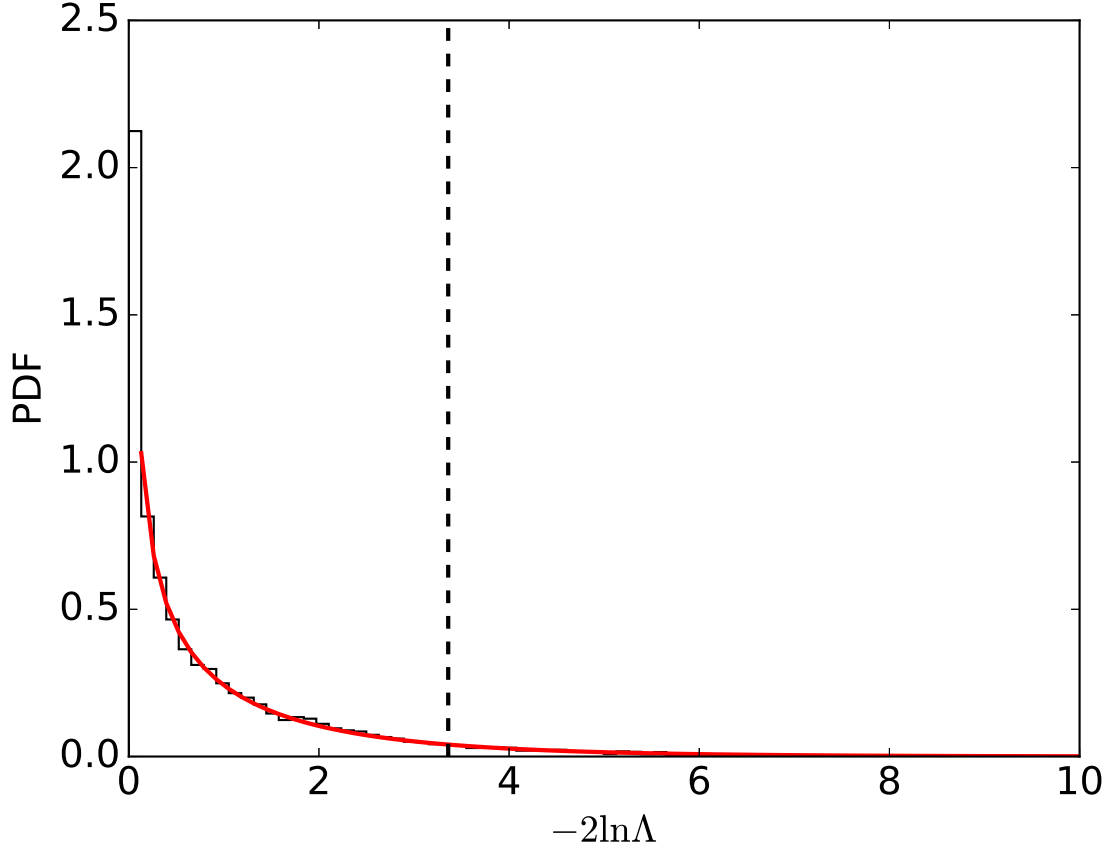


Figure 2: The test statistic for $H_0 = 150$, $H_1 \neq 150$. The prediction from Wilk's Theorem is overlaid. The sample data set's value of $-2 \ln \Lambda$ is denoted by the dashed line.

- H_0 : $\mu = 150$
- H_1 : $\mu \neq 150$

We repeat the analysis from the previous section, and the position of the same sample of 10 from the original $\mu = 100$ distribution is shown by the dashed line. The p-value of this observation is 0.0634, suggesting this result is still not significant at a 5% level.

(c)

Wilk's theorem states that in the large sample size limit that $-2 \ln \Lambda$ asymptotically tends to a χ^2 distribution with degrees of freedom equal to the difference in dimensionality of the hypotheses. In both cases we have studied, the null hypothesis is a point of dimension zero, and the alternative hypothesis is the real number line with a hole. Thus, we expect the $-2 \ln \Lambda$ to follow a χ^2 distribution with 1 degree of freedom. In Figures 1 and 2, this is shown by a red line. The p-values predicted for the hypotheses are 0.3822 and 0.0668, respectively computed from the χ^2 CDF. This shows very good agreement despite only having 10 samples.

Problem 2

(a)

We are asked to find the properties of the 90% confidence interval on the number of counts expected in a Poisson-distributed measurement. The number observed, $n = \hat{\mu}$, is the maximum likelihood estimator for μ , the expected number of counts in each bin. The confidence interval $[a, b]$ is in general determined implicitly from the equations [Cowan, 1998],

$$\alpha = \int_{\hat{\mu}}^{\infty} f(\hat{\mu}|a) d\hat{\mu} \quad (3)$$

$$\beta = \int_{-\infty}^{\hat{\mu}} f(\hat{\mu}|b) d\hat{\mu} \quad (4)$$

where α and β are the right and left significance probabilities, and $f(\hat{\mu})$ is the PDF of the estimator. Following the argument in Cowan [1998], we evaluate these integrals in semi-analytic fashion in terms of known functions. Starting with the first. and since the Poisson distribution is discrete, we begin with,

$$\alpha = 1 - \sum_{k=0}^{\hat{\mu}-1} \frac{a^k}{k!} e^{-a} \quad (5)$$

$$\geq 1 - \int_{2a}^{\infty} g_{\chi^2}(\psi | df = 2\hat{\mu}) d\psi \quad (6)$$

$$= 1 - G_{\chi^2}(2a | df = 2\hat{\mu}) \quad (7)$$

where g_{χ^2} and G_{χ^2} are the χ^2 PDF and CDF respectively. The second step follows from the fact that each term in the Poisson distribution has the same form as the χ^2 PDF, and as the mean becomes large, the sum approaches the integral. Another interesting point that stems from this is the fact that the interval will be overcovered. This simply arises from how the χ^2 CDF bounds the Poisson sum. We can invert and solve for a trivially as,

$$a = \frac{1}{2} G_{\chi^2}^{-1}(\alpha | df = 2\hat{\mu}). \quad (8)$$

We can repeat this logic for the β sum,

$$1 - \beta = 1 - \sum_{k=0}^{\hat{\mu}} \frac{b^k}{k!} e^{-b} \quad (9)$$

$$\leq 1 - \int_{2b}^{\infty} g_{\chi^2}(\psi | df = 2(\hat{\mu} + 1)) d\psi \quad (10)$$

$$= 1 - G_{\chi^2}(2b | df = 2(\hat{\mu} + 1)), \quad (11)$$

giving,

$$b = \frac{1}{2} G_{\chi^2}^{-1}(1 - \beta | df = 2(\hat{\mu} + 1)). \quad (12)$$

We now wish to consider three intervals of confidence level, $1 - \gamma = 0.9$. The first interval is symmetric with $\alpha = \beta = 0.5\gamma$. The other two, the upper and lower limit intervals, take either $\alpha = 0.1$ or $\beta = 0.1$ respectively. From Eqs. (8) and (12), we determine that for the symmetric interval,

$$[a_{sym}, b_{sym}] = \left[\frac{1}{2} G_{\chi^2}^{-1}(0.05 | df = 2\hat{\mu}), \frac{1}{2} G_{\chi^2}^{-1}(0.95 | df = 2(\hat{\mu} + 1)) \right], \quad (13)$$

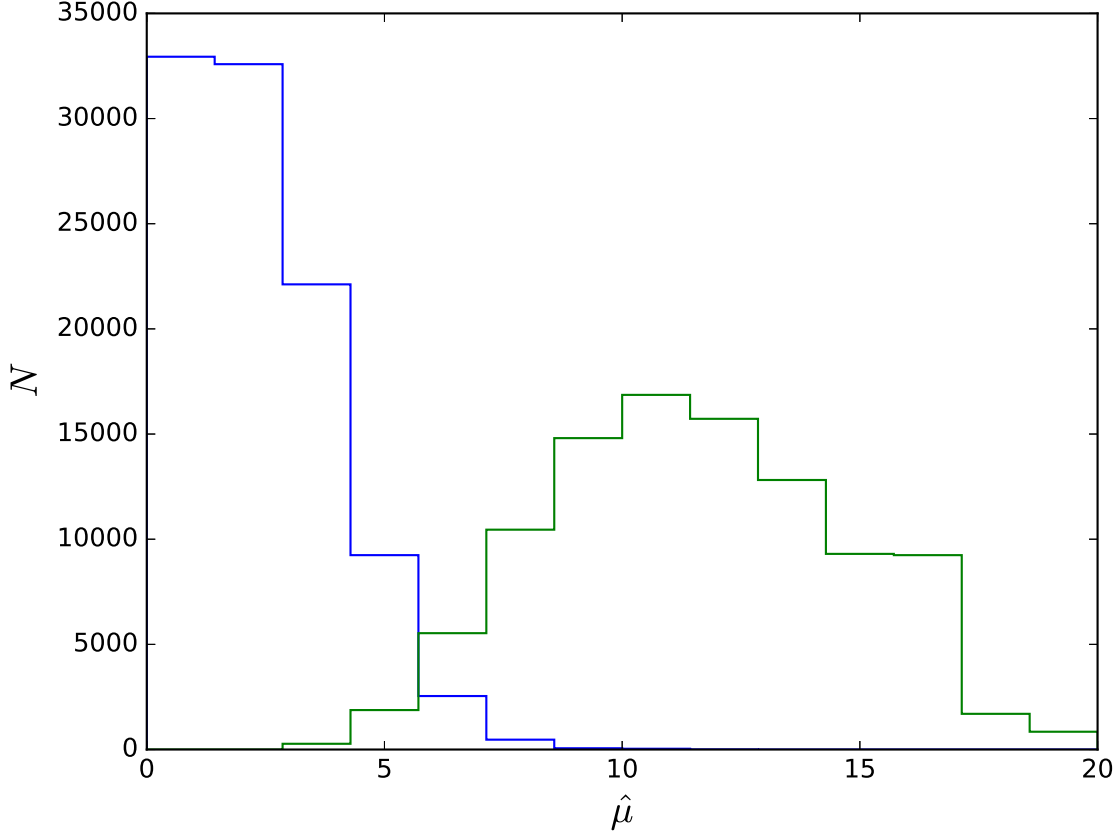


Figure 3: The distribution of b (blue) and a (green) computed from Eqs. (8) and (12) for a symmetric confidence interval and 10^5 draws.

for the upper limit interval,

$$b_{up} = G_{\chi^2}^{-1}(0.9 | df = 2(\hat{\mu} + 1)), \quad (14)$$

and for the lower limit interval,

$$a_{low} = \frac{1}{2} G_{\chi^2}^{-1}(0.1 | df = 2\hat{\mu}). \quad (15)$$

(b)

To test the validity of Eqs. (8) and (12), we draw 10^5 samples from a Poisson distribution of mean 5.7. For each of these we compute the upper and lower limits for the symmetric and limiting intervals at $\gamma = 0.1$. Figure 3 shows the distribution of a and b for the symmetric interval. The results are broadly in agreement with the tabulated values in Cowan [1998]. Similarly, for the upper and lower one-sided intervals, we make the same plot in Figure 4. Again, these intervals broadly agree with tabulated values for a Poisson distribution.

An interesting point to make is that the confidence intervals actually cover $\gamma_{sym} = 0.94736$, $\gamma_{low} = 0.93520$, and $\gamma_{up} = 0.92319$. The fact that we have overcovered is a known effect that occurs due to the discrete nature of the Poisson distribution. When we use the χ^2 CDF to compute these quantities, we are using an upper bound on the value of the sum. The discrepancy disappears if one takes higher μ , since the relative spacing between significant points in the CDF becomes small. That is to say, the CDF itself becomes broader, and a $\Delta\mu$ of 1 is no longer as large.

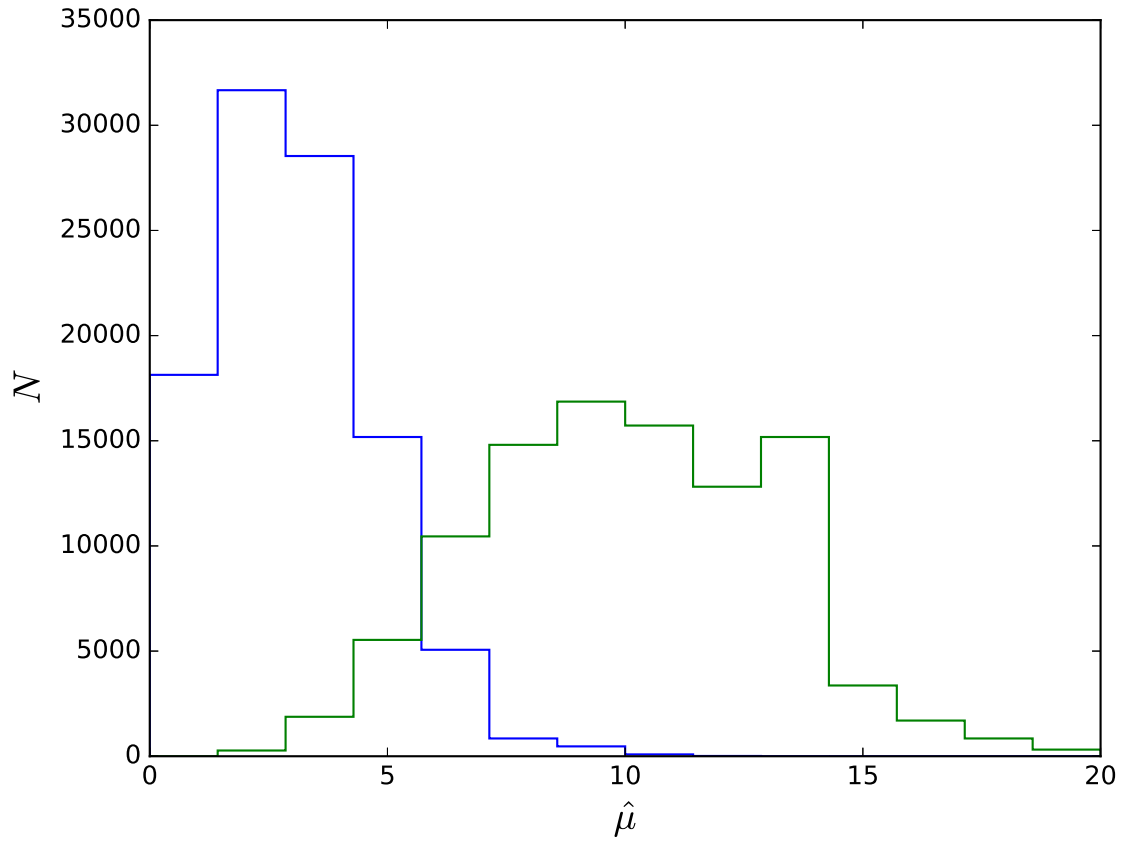


Figure 4: The distribution of b (blue) and a (green) computed from Eqs. (8) and (12) for the upper limit and lower limit intervals and 10^5 draws.

Another interesting remark is the case where $n = 0$. This can happen quite easily if the mean is small, but only a minor modification to the calculation is needed. In this case, we are only able to establish an upper bound, which we do at a level β . The calculation for b is the same, and we take $a = 0$.

References

G. Cowan. *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. ISBN 9780198501558. URL <https://books.google.ca/books?id=ff8ZyW0n1JAC>.