STA500 Introduction to Probability and Statistics 2, autumn 2018.

Exercise set 5

Exercise 1:

Assume that alarms to a fire station arrive as a Poisson process. Based on the information about the times between such alarms, we want to estimate the parameter λ, the expected number of alarms per hour, in the Poisson process. We know that the times between alarms, $T_1, \ldots, T_n$ will be exponentially distributed with expectation $1/\lambda$. I.e. the probability density is given as:

$$f(t) = \lambda e^{-\lambda t} \quad , t \geq 0$$

a) Show that the maximum likelihood estimator (MLE) for λ becomes $\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} T_i}$. Calculate the estimate when the following 12 times between alarms have been recorded: 10.3, 5.5, 18.0, 12.1, 9.1, 1.2, 5.6, 18.6, 7.7, 3.9, 9.9 and 6.6.

b) Use results on transformations and linear combinations to explain that $Z = 2\lambda \sum_{i=1}^{n} T_i$ has a χ^2_{2n} distribution.

It can be shown that if $Z \sim \chi^2_{2n}$ then $\mathrm{E}(\frac{1}{Z}) = \frac{1}{2(n-1)}$

c) Use the results above to calculate $\mathrm{E}(\hat{\lambda})$.
Is $\hat{\lambda}$ unbiased? If not, can you suggest a modified estimator which will be unbiased? Comment these results and the fact the $\hat{\lambda}$ is an MLE.

d) Use the MLE estimator $\hat{\lambda}$ and the results from 1b) as starting point, and derive a 95% confidence interval for λ. Calculate the interval numerically for the measurement given in point 1a).
(Hint: $Z = 2\lambda \sum_{i=1}^{n} T_i = 2\lambda n \sum_{i=1}^{n} T_i/n = 2\lambda n/\hat{\lambda}$)

Exercise 2:

Let $X_1, X_2, \ldots, X_n$ be independent and exponentially distributed with expectation β. I.e

$$f(x) = \frac{1}{\beta} e^{-x/\beta} \quad , x \geq 0,$$

a) Combine results on transformations of variables and results on sums of variables from the collection of formulas to find the distribution of $Z = \sum_{i=1}^{n} 2X_i/\beta$.
(Hint, start by considering the distribution of $2X_i/\beta$.)

b) Use results on transformations of variables to find the distribution of $V = \sum_{i=1}^{n} X_i$.
Which result from the lectures have we shown now?

Exercise 3:

In an opinion poll the proportion of voters voting for different parties are examined. We shall in this exercise consider the proportion of voter that would vote for Ap (the Norwegian labour party), and call this proportion p. Assume that the opinion poll is carried out by asking 1000 randomly selected voters. Let X = the number of voters that would vote for Ap.

a) Explain which exact distribution X is having, and explain why this distribution can be approximate by a binomial distribution with parameters $n = 1000$ and p.

b) If $p = 0.275$, what is the probability that X becomes larger than 300?

In practise p is unknown, the purpose of the the opinion poll is to estimate p.

c) Show that the maximum likelihood estimator (MLE) for p becomes $\hat{p} = \frac{X}{n}$. Calculate $\mathrm{E}(\hat{p})$ and $\mathrm{Var}(\hat{p})$.

Of the 1000 voters asked, 297 reported that they would vote Ap.

d) Derive an approximate 95% confidence interval for p, and calculate the interval numerically for the observation above.
(Hint: Use the central limit theorem as starting point. Further you can in the derivation use that $p(1-p) \approx \hat{p}(1-\hat{p})$ when n is large.)

Some answers:

1: a) 0.11, c) $\lambda\frac{n}{n-1}$, biased estimator: $\frac{n-1}{\sum_{i=1}^{n} T_i}$, d) [0.06, 0.18]
2: b) Gamma distribution with $\alpha = n$ and $\beta = \beta$;
3: b) 0.0351; d) [0.269, 0.325] ;