STAT 135 Lab 4

GSI: Yutong Wang

Feb 14, 2020

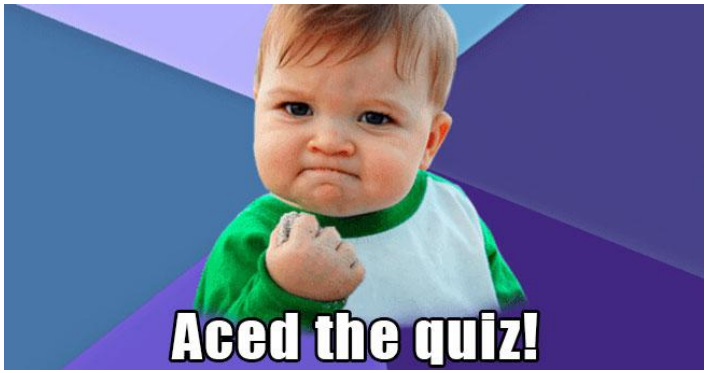
Berkeley
UNIVERSITY OF CALIFORNIA

To-do Today

1. Quiz 1 recap
2. Order statistics
3. Fisher information and Cramer-Rao inequality
4. Mean squared error
5. MSE and MoM estimation
6. ggplot in R (dplyr)

Source: Facebook group: statistical statistics memes

Quiz 1 Recap



1. Problem-solving skills:
   - What is the question asking for?
   - What information is given?
   - Make sure you follow each step and the definition!

Problem 1.

(a) $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} Beta(\alpha, \alpha)$. Give the MOM estimator $\hat{\alpha}$ of the parameter α.

(b) Given $n = 8$, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = 4$, calculate your $\hat{\alpha}$.

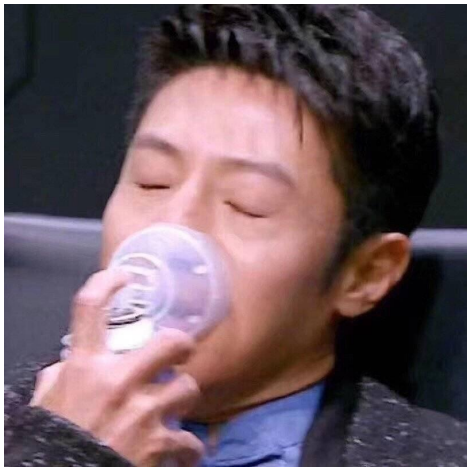
- Confusion over concepts:

$$\mathbb{E}[X^2] = \mathrm{Var}(X) + \mathbb{E}^2(X)$$

$$\hat{\mu_2} \overset{?}{=} s^2 + (\bar{X})^2$$

Quiz 1 Recap

2. Don't be afraid of the wording of the problem. They are just STATISTICS. ^



Problem 3.

During the office hour, a GSI is counting the time gap between students who come. Each time he can only talk with one student. He writes down the time of ith student's arrival and denotes it as T_i, and then records the $(i+1)$th student's arrival time as T_{i+1}, and then he calculates $\Delta T_i = T_{i+1} - T_i$. The unit of ΔT_i is minute. Assuming that $\Delta T_1, \Delta T_2, ..., \Delta T_n \overset{i.i.d.}{\sim} Exp(\lambda)$, and after some counting and recording, he estimated a 95% confidence interval for $1/\lambda$ as $[0.836, 1.164]$.

Facts and hints:

1. He used $\overline{\Delta T}$ to estimate $1/\lambda$
2. If the arrival time gaps *i.i.d.* follow $Exp(\lambda)$, then given any time t, the number of arrivals within t follows $Pois(\lambda t)$.

Questions:

(a) What's his value of estimator for $1/\lambda$? What's the value of estimator for λ?

(b) How many time gaps did he observe? (i.e. find n).

(c) If this GSI uses his observation to predict how crowded it will be for tomorrow's office hour, what's the expectation and variance of the number of students who come within 10 minutes at the beginning of tomorrow's office hour?

Order Statistics (Rice 8.10.27)

Suppose that certain electronic components have lifetimes that are exponentially distributed: $f(\tau|t) = \frac{1}{\tau}e^{-\frac{1}{\tau}}$ for $t \geqslant 0$. Five new components are put on test, the first one fails at 100 days, and no further. observations are recorded.

(a) What is the likelihood function of τ?

(b) What is the MLE of τ?

(c) What is the sampling distribution of the MLE?

(d) What is the standard error of the MLE?

Hint: The minimum of independent exponentials is itself an exponential. Why?

Source: Matrix 4

Fisher Information and Cramer-Rao Inequality (Rice 8.10.19)

Suppose that $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$.

(a) If μ is known, what is. the MLE of σ?

(b) If σ is known, what is the MLE of μ?

(c) In part (b), does any other *unbiased* estimate of μ have smaller variance?

Mean Square Error

If the random variable Y has a Binomial(n, p) distribution (with n known), consider the two estimators

$$\hat{p}_1 = \frac{Y}{n} \qquad\qquad \hat{p}_2 = \frac{Y+1}{n+2}$$

(a) Show that $\hat{p}_1$ is unbiased.

(b) What is the bias of $\hat{p}_2$?

(c) Derive MSE$(\hat{p}_1)$ and MSE$(\hat{p}_2)$. *Hint:* Recall that $\text{MSE}(\hat{\theta}) = [\text{Bias}(\hat{\theta})]^2 + Var(\hat{\theta})$

(d) For what values of p is MSE$(\hat{p}_1) <$ MSE$(\hat{p}_2)$?

MSE and MoM Estimation (Rice 8.10.47)

The Pareto distribution has been used in economics as a model for a density function with a slowly decaying tail:

$$f(x|x_0,\theta) = \theta x_0^{\theta} x^{-\theta-1}, \qquad x \geqslant x_0, \theta > 1$$

Assume that $x_0 > 0$ is given and that $X_1, \ldots, X_n$ is an i.i.d. sample.

(a) Find the MoM estimate of θ.

(b) Find the MLE of θ.

(c) Find the asymptotic variance of the MLE.

WHAT IS GGPLOT AFTERALL???

How mornings look like for most people:

```
me %>%
  wake_up() %>%
  get_out_of_bed() %>%
  get_dressed() %>%
  leave_house()
```

How my mornings look like most of the time:

```
leave_house(get_dressed(get_out_of_bed(wake_up(me))))
```

Source: Facebook group: R memes for statistical friends

Useful resources for learning R

1. **What do the code chunks mean?**
   - R Markdown from R studio: https://rmarkdown.rstudio.com/lesson-3.html
2. **How to code gracefully?**
   - Google's R Style Guide: https://google.github.io/styleguide/Rguide.html
3. **How to make fancy ggplot?**
   - R open source graphing library: https://plot.ly/r/
   - A tour through the visualization zoo: https://queue.acm.org/detail.cfm?id=1805128
4. **How to make color choices?**
   - Color Brewer: http://colorbrewer2.org/?#type=sequential&scheme=BuGn&n=3
   - Coolors.do: https://coolors.co/bf4e30-c6ccb2-093824-e5eafa-78fecf
5. **Is there any overview of different data structures in R?**
   - R for Data Sciences: https://r4ds.had.co.nz/