

STATS 412

Fourteenth Class Note

In Son Zeng

04 November, 2018

My Office Hour:

The office hours this week are on **16:30 - 18:00 Tuesday** and **13:30 - 15:00 Friday**, at **USB 2165**. You may check the campus map to get to my office. I am prepared for your questions, so please feel free to come to my office hours.

Calculus Review:

• To compute the Maximum Likelihood Estimator (MLE), you may encounter the difficulty for partial differentiation. If you have studied MATH 215 or the equivalent class before, you may review the notes. If you do not know how to perform partial differentiation, you may refer to the following websites for reference:

• <http://tutorial.math.lamar.edu/Classes/CalcIII/PartialDerivsIntro.aspx>

• <https://www.khanacademy.org/math/multivariable-calculus/multivariable-derivatives/partial-derivative-and-gradient-article/a/introduction-to-partial-derivatives>

These are great practices to prepare you with essential calculus skills and knowledge of distributions for the subsequent homework and the exam 2.

Homework Grading Policy:

Please include the final answer for each homework question. If the final answer is not included, you will risk 0.5 points for each missing part.

Homework 7 Short Version Comments:

• Make sure you are applying the correct formula for the MSE: $MSE(x) = [Bias(x)]^2 + Var(x)$. Additionally, there is always a tradeoff between increasing the bias while reducing the variance, or decreasing the bias while increasing the variance. In the realm of statistics, depending on the context, we may select an estimator which minimizes the variance, or is unbiased, or minimizes the MSE.

• Bias can be either positive, 0 or negative. If the bias is positive, it means the estimator overestimates the parameter. If the bias is 0, it means the estimator is unbiased. If the bias is negative, it means the estimator underestimates the parameter.

• In question 2 and 3, a majority of submissions show no work in checking that the MLE is indeed the maximum by taking the second derivative. This is a big problem since based on calculus, when the first derivative of a parameter equals to 0, then the parameter achieves **local maximum** if the second derivative (curvature) is negative, or the parameter achieves **local minimum** if the second derivative is positive.

• For question 4 - 8, a vast majority of students mistakenly did two things. The first one is the distribution of the sample mean \bar{X} and the summation of observations S_n should be approximately normal, but NOT exactly normal. As a result, the second thing, the probability should be approximately equal, but not equal to the probability given by the z-score. That is the trickiest part of CLT which worth practices. See Homework 7 comments and the solution and come to my office hours for detail.

Key Points during Lecture:

Test Reminder: There will be the standard normal cumulative distribution function table provided for reference during the test. This can save your time in finding the correct probability or z-score to answer the problems such as: (1) Central Limit Theorem related approximation problems, (2) Normal distribution problems, (3) Deriving p-values for hypothesis testing.

Large Sample Size: Regarding the notion of large sample size, in our class we use $n \geq 30$ as criteria.

p-value: We officially introduced the definition of p-value as follows: A p-value is the (conditional) of observing a test statistics as extreme (or weird) as or more extreme than what we observed, assuming that the null hypothesis (H_0 , the hypothesis that we would like to find evidence to reject) is true.

Report p-values It is always great to report the p-values because everyone may define a small p-value differently. Usually, people choose either $\alpha = 0.05$ or $\alpha = 0.01$. This “selected” level is called **significance level**.

Story for Student's t distribution: The Student's-t distribution was published around 1908, when a statistician William Gossett thought that their company products generated some samples with a distribution very closed to the normal distribution, but with heavier tails. To address the difference, the Student's t distribution is constructed with the additional concept of **degree of freedom** to capture the discrepancy of the extra variability in the sample.

More detail: <https://onlinelibrary.wiley.com/doi/full/10.1002/cem.2713>

How to calculate the quantile or the probability in t-distribution: Student's t distribution introduces the concept **degree of freedom** so that the R code for the t distribution involves assigning the degree of freedom. The R-codes to compute the probability of $P(t < a)$ with sample size n is given by the format `pt(a, n)`. The R-codes to compute the $(100 \times a)^{th}$ quantile with sample size n is given by the format `qt(a, n)`.

```
pt(2,200) # The probability of  $P(t < 2)$  if the sample size is 200.
```

```
## [1] 0.9765734
```

```
qt(0.95,2000) # The 95-th quantile of Student's t distribution when sample size is 2000
```

```
## [1] 1.645616
```

Steps of Hypothesis Testing: Will be added next time. You may refer the notes first.

Last Comment:

Please inform me to fix the typos and grammatical mistakes if they exist. It is a great practice of writing and I appreciate your help!