STATS 412

Twelveth Class Note

In Son Zeng

26 October, 2018

My Office Hour:

My office hours 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. During the fall break, I have extra office hours on **14:00 - 16:00 Monday** and **16:30 - 18:00 Tuesday** before the midterm.

Calculus Review:

• After the exam 1, the requirements for integration will be lowered. However, 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-articles/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.

Key Points during Lecture:

Poisson Distribution Review: It is clear in the lecture that Poisson distribution is an important and special distribution because its mean and variance are the same. That is, for a Poisson random variable $X \sim Poisson(\lambda)$:

$$\boxed{E(\bar{X}) = E(X) = \lambda, Var(\bar{X}) = \frac{\lambda}{n}, E(X) = Var(X) = \lambda} \tag{1}$$

If $X_1, X_2, ...X_n$ are independent Poisson random variable of $Poisson(\lambda)$, then we have the joint pmf as:

$$\boxed{p(x_1, x_2, ..., x_n) = p(x_1) \cdot p(x_2) \cdot ...... \cdot p(x_n) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}} \tag{2}$$

Maximum Likelihood Estimate (MLE): As taught in the last lecture, in the Poisson case, from the independence of $X_1, X_2, ...X_{30}$ we use the multiplicative rule to set up the **likelihood function**. The likelihood function is in the form:

$$\boxed{p(X_1 = x_1) \cdot p(X_2 = x_2) = ...p(X_{30} = x_{30}) = \prod_{i=1}^{30} \frac{e^{-\lambda} \cdot \lambda^{x_i}}{x_i!} = \frac{e^{-30\lambda}\lambda^{\sum_{i=1}^{30} x_i}}{\prod_{i=1}^{30} x_i!}} \tag{3}$$

The MLE can be obtained by finding the value of λ which maximizes this likelihood function. As taught in the lecture, we may take the derivative of the likelihood function with respect to λ.

However, this is not the whole story. In many case, it is difficult to take the derivative for a complicated likelihood function due to the complexity of probability distribution. One of the remedy is to take log to the likelihood function; such function is widely known as log-likelihood. Since log is a monotone increasing function, we can take the derivative of the log-likehood function to obtain the same MLE result.

• Maximum Likelihood Estimate (MLE) tends to be harder from the past experiences. Please ask questions during the office hours whenever you encounter trouble doing homework or revising the notes. I will include an extra question about MLE before the exam for revision. Again, I hope that students can help each other and post possible solution for the extra problems and I will definitely check.

Estimator and Estimate

• In statistics, we should be conscious about the differences between an estimator and an estimate. Generally, an estimator is a mathematical function to produce an estimate, while an estimate is a computed value based on data which it may be modeled as random variables. This notion makes the estimate into a random variable and the estimator into a function of the random variable.

• In the book Introduction to Econometrics, the author James H. Stock, Mark W. Watson claim that an estimator is a function of a sample of data that to be drawn randomly from a population such that it gives an educated guess of the value from the population. An estimate is the numerical value of the estimator when it is calculated using data from a specific sample.

For more explanation, you are encouraged to look that the following posts:

(1) https://stats.stackexchange.com/questions/7581/what-is-the-relation-between-estimator-and-estimate

(2) https://en.wikipedia.org/wiki/Estimator

• In some cases, we may have different estimators producing the same value of the estimate.

• During the class we are given some examples. $\bar{X}$ is an estimator for μ, while $\bar{x}$ is the estimate because it is derived from a sample. Estimate is the realized, observed value from data, while the estimator is a function.

MLE for normal distribution: If, the joint PDF (likelihood function) is $f(x_1, x_2, ..., x_n; \mu, \sigma^2) = (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} \cdot exp\Big(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\Big)$. Take the log we obtain the log-likelihood function:

$$\boxed{ln(f(x_1, x_2, ..., x_n; \mu, \sigma^2) = -\frac{n}{2}log(2\pi) - \frac{n}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \tag{4}$$

To find out the MLE estimator for the mean, we take the first derivative with respect to μ (I skipped the verification part which requires you to take the second derivative of μ and conclude that the second derivative of the log-likelihood function is negative for all x), and set the derivative as 0 (if you are having trouble this step, please also revise the chain rule in calculus:

$$\boxed{\frac{\partial ln(f(x_1, x_2, ..., x_n; \mu, \sigma^2)}{\partial \mu} = -\frac{1}{2\sigma^2}\cdot 2 \cdot (-1) \cdot \sum_{i=1}^{n}(x_i-\mu) =^{set} 0 \rightarrow \sum_{i=1}^{n} -n\mu = 0 \rightarrow \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}} \tag{5}$$

To find out the MLE estimator for the variance, we take the first derivative with respect to σ^2 (Again, I skipped the verification part which requires you to take the second derivative of σ^2 and find the result $\frac{\partial^2 ln()}{\partial(\sigma^2)^2} = -\frac{n}{2}$ that the second derivative of the log-likelihood function is negative for all x), and set the derivative as 0 (if you are having trouble this step, please also revise the chain rule in calculus:

$$\boxed{\frac{\partial ln(f(x_1, x_2, ..., x_n; \mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\cdot \sum_{i=1}^{n}(x_i-\mu)^2 =^{set} 0 \rightarrow \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 = n \rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}} \tag{6}$$

I added a hat on the top of the μ and σ^2 to distinguish that they are, in fact, an estimator, In addition, for the variance, if we consider the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, we the MLE of the variance can be rewritten as $\hat{\sigma}^2 = \frac{n-1}{n}S^2$.

Properties of MLE:

• Since the MLE does not necessarily exist and unique, the derivation of MLE may not work.

• Invariance Principle: For example, $\hat{\sigma}^2 = \frac{n-1}{n}S^2$ implies $\hat{\sigma} = \sqrt{\frac{n-1}{n}S^2}$.

• Since MLE approximates the minimum variance (MVUE), it does not necessary achieve the minimum variance. Students do not need to revise whether MLE achieves the minumum variance for the exam 2.

Gamma distribution and its Family:

The probability density function of gamma(α, β) distribution is given by:

$$\boxed{f(x; \alpha, \beta) = \begin{cases} \frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)}, \ \alpha > 0, \ \beta > 0, \ 0 < x < \infty \\ 0, \ otherwise \end{cases}} \tag{7}$$

Then, its expectation and variance are derived by:

$$\boxed{E(X) = \alpha \cdot \beta} \tag{8}$$

$$\boxed{V(X) = \alpha \cdot \beta^2} \tag{9}$$

From the lecture, we may have trouble in performing integrating by part for

$$\boxed{f(x) = \begin{cases} \frac{9}{4}xe^{-\frac{3x}{2}}, 0 < x < \infty \\ 0, \ otherwise \end{cases}} \tag{10}$$

In this case, to compare the coefficients, we find that $\alpha - 1 = 1 \rightarrow \alpha = 2$ and $\frac{1}{\beta} = \frac{3}{2}$. Therefore, we conclude that it is $Gamma(2, \frac{2}{3})$ distribution. Hence, the expectation and variance of this case are:

$$\boxed{E(X) = \alpha \cdot \beta = 2 \cdot \frac{2}{3} = \frac{4}{3}} \tag{11}$$

$$\boxed{V(X) = \alpha \cdot \beta^2 = 2 \cdot \left(\frac{2}{3}\right)^2 = \frac{8}{9}} \tag{12}$$

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!