STATS 412

Seventh Class Note

In Son Zeng

01 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.

Calculus Review:

• I saw a great suggestion from a student who mentioned that there is a great resource for self-study about the calculus. The website is: https://instruct.math.lsa.umich.edu/lecturedemos/ma215/docs/

• I also uploaded a book for probability to the Piazza. Particularly, I recommend you to read (Type the page number in PDF):

(1) Improper integral: page 203 - 210

(2) Double integral: page 214 - 218

(3) Cumulative distribution function: page 224 - 229

(4) Expectation and Variance: page 235 - 244

(5) Joint Distributions: page 299 - 304

These are great practices to prepare you with essential integration skills for the subsequent homework and the midterm.

Reminders for Assignment 4

Key Points during Lecture:

Variance: The variance of $X_1 + X_2 + X_3$ is generally different from the variance of $3X$. The reason is:

$$\boxed{Var(X_1 + X_2 + X_3) = Var(X_1) + Var(X_2) + Var(X_3) + 2cov(X_1, X_2) + 2cov(X_1, X_3) + 2cov(X_2, X_3)} \tag{1}$$

$$\boxed{Var(3X) = 3^2 Var(X) = 9Var(X)} \tag{2}$$

Test Reminder: Do not lose point by not writing the 0 otherwise, whenever you encounter a question asking you to write down a probability distribution.

Conditional Probability: During the previous class there is a great question about conditional probability: given $f(x, y) = 0.000006e^{-0.001x-0.002y}$ for $0 < x < y < \infty$ (0 otherwise), derive $P(Y > 2000|X > 1500)$. From the equation list, you may know

$$\boxed{P(Y > 2000|X > 1500) = \frac{P(X > 1500 \cap Y > 2000)}{P(X > 1500)} = \frac{\int\int f(x, y)dydx}{\int_{1500}^{\infty} f_x(x)dx}} \tag{3}$$

How can we derive the final answer? After calculation, you may refer the sixth class note for solution.

Property of Independence: Recall if random variables $X_1, X_2, ..., X_n$ are independent, then (1) If $X_1, X_2, ..., X_n$ are jointly discrete/continuous, the joint pmf/pdf is equal to the product of the marginals.

(2) If X, Y are independent and (a) jointly discrete, and given respectively the marginal of x $p_X(x) > 0$ and marginal of y $p_Y(y) > 0$, (b) jointly continuous, and given respectively the marginal of x $f_X(x) > 0$ and marginal of y $f_Y(y) > 0$, then by making comparison of terms we have:

$$\boxed{p_{Y|X}(y|x) = p_Y(y), \quad p_{X|Y}(x|y) = p_X(x)} \tag{4}$$

$$\boxed{f_{Y|X}(y|x) = f_Y(y), \quad f_{X|Y}(x|y) = f_X(x)} \tag{5}$$

Check Independence: For discrete random variables X and Y, to check independence, we basically need to check every pair of $p(x) \cdot p(y) = p(x, y)$. There are tricky cases when the result holds true for some pairs but not all pairs. Now we summarize the relationship between independence and correlated. If random variables X, Y are independent, then $cov(X, Y) = 0$. However, if $cov(X, Y) = 0$, X, Y are not necessarily independent. In short, we have X, Y *independent* $\rightarrow cov(X, Y) = 0$, but $cov(X, Y) = 0 \nrightarrow X, Y$ *independent*.

Measurement Error Example: The four readings (in pounds) are 148, 151, 150, and 152. Each time the person gets off the scale, the reading is 2 pounds. We can estimate the uncertainty by calculating the standard deviation. Let us break it down.

$$\boxed{E(X) = \frac{148 + 151 + 150 + 152}{4} = 150.25} \tag{6}$$

$$\boxed{Var(X) = \frac{1}{4-1}\Big((-2.25)^2 + 0.75^2 + (-0.25)^2 + 1.75^2\Big) = 2.9167} \tag{7}$$

$$\boxed{SD(X) = \sqrt{Var(X)} = 1.7078 \approx 1.71} \tag{8}$$

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!