STAT 2507

Tutorial 3

Contact Information

Email: melissavanbussel@cmail.carleton.ca

GitHub: www.github.com/melissavanbussel

Section A3: Mondays 4:35-5:25

Section C2: Tuesdays 2:35-3:25

Section A5: Thursdays 12:35-1:25

Section A8: Thursdays 1:35-2:25

Today's Tutorial

- Feedback from previous tutorials
- Briefly: 3 common mistakes from Test 1
- Quick review: What is a CDF?
- Using SPSS to generate the CDF of a Binomial and a Poisson random variable

Feedback from previous tutorials

- THANK YOU for your feedback, it is very helpful!
- I will try to leave more time at the end for assignment questions since some students can't stick around after the tutorial -- let's give priority to those who have to leave right at the end of tutorial
- Things we will continue doing: common mistakes from most recently marked assessments, quick review of the core concepts before SPSS
- Unfortunately, I'm not allowed to use Zoom for the tutorials. We have to continue using BigBlueButton

Types of Variables

- A variable can either be qualitative, or quantitative.
- If a variable is quantitative, it can be further classified as either continuous or discrete.
- A qualitative variable **cannot** be discrete or continuous -- these words only apply to quantitative variables!

Finding the least-squares regression line

- You need to use the formulae provided on the formula sheet
- You can’t just pick 2 arbitrary points and connect them!

Expected Value and Variance

- Sometimes, you will hear people refer to the expected value as the “mean”
- There is a **difference** between finding the mean and variance of a sample of data, and finding the mean and variance of a random variable.

Find the mean and variance: sample of data

- I decide to take a sample of students from our course. The heights of the student in this sample are (in centimetres): 158, 188, 172, 170.
- Mean = 172, Variance = 152

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Find the mean and variance: random variable

- I roll an unfair die. Let X denote the number that I roll. X has the following probability distribution:

X	1	2	3	4	5	6
P(X = x)	0.15	0.1	0.4	0.05	0.2	0.1

- Expected value of X = $1(0.15) + 2(0.1) + 3(0.4) + 4(0.05) + 5(0.2) + 6(0.1) = 3.35$
- Variance of X = $1^2(0.15) + 2^2(0.1) + 3^2(0.4) + 4^2(0.05) + 5^2(0.2) + 6^2(0.1) - 3.35^2 = 2.3275$

$$E(X) = \sum_i x_i \, p(x_i)$$

$$V(X) = \sum_i x_i^2 p(x_i) - \left(\sum_i x_i p(x_i)\right)^2$$

Review: What's a CDF?

- **C**umulative **D**istribution **F**unction
- It's the probability that the random variable takes a value less than or equal to what's specified; i.e., F(x) = P(X <= x)
- Starts at 0, goes up to 1
- Consider the previous example

X	1	2	3	4	5	6
P(X = x)	0.15	0.1	0.4	0.05	0.2	0.1
F(x)	0.15	0.25	0.65	0.70	0.90	1

What would F(0) be? What about F(7)?