1️⃣

Introduction

Status	Done
Type	Lecture
Course	Mathematical Simulation
Found a mistake?	fadikatrina@me.com

Quick General Statistics Background Primer
- “Classical” Variance
- Standard Deviation
- Covariance
- Correlation

Basics of discrete random variables
- Expected Value
- Variance

Variance Sum Law
Variance of an Average
Central Limit Theorem
Monte Carlo Integration

Read the intro from the MM notes, most of it applies here as well.

Quick General Statistics Background Primer

“Classical” Variance

> https://www.youtube.com/watch?v=x0rmUXWtSS8, https://www.youtube.com/watch?v=SzZ6GpcfoQY

- “The average of the squared differences from the mean”, notation σ^2

- Used to see how “spread out” the data is from the mean
- Yields slightly weird results because all the values are squared

$$\frac{\sum(datapoint - mean)^2}{n}$$

- or $n - 1$ if its a population (Bessel’s correction)

Standard Deviation

- “Extension of variance” but slightly better, because the result is more interpretable for our use case, notation σ
- We simple square root the variance $\sqrt{variance}$

Covariance

> https://www.youtube.com/watch?v=qtaqvPAeEJY

- We use covariance to try to determine how related are two variables

$$\frac{\sum(x - \bar{x})(y - \bar{y})}{n - 1}$$

- Note: if you find the covariance of a variable with itself (x and y the same variable) you calculate the variance of that variable
- Problem: its sensitive to the values/magnitudes of the scales, this is solved with correlation

Correlation

> https://www.youtube.com/watch?v=xZ_z8KWkhXE

- Normalised covariance, deals with the issue that covariance depends on the scales

$$\frac{Covaraince(x, y)}{\sqrt{Variance(x)} \cdot \sqrt{Varaince(y)}}$$

Basics of discrete random variables

> http://www.stat.yale.edu/Courses/1997-98/101/ranvar.htm

- A ***discrete random variable*** is one which may take on only a countable number of distinct values such as 0,1,2,3,4
- Examples
  - the number of children in a family
  - the Friday night attendance at a cinema
  - the number of patients in a doctor's surgery
  - the number of defective light bulbs in a box of ten.
- **Probability distribution/probability function/probability mass function**: list of probability of each possible value
  - Simple hypothetical example for number of children in a family in a certain population:

x	0	1	2
p(x)	0.16	0.48	0.36

Expected Value

> https://www.youtube.com/watch?v=Vyk8HQOckIE

- Expected value is the theoretical mean, it does not mean the most likely value, and a lot of the time its not even possible for the random variable to have this value, the way it is named is unfortunate, notation:

$$E(X) \ or \ \mu$$

- How its calculated:

$$E(X) = \sum x \cdot p(x)$$

- So multiplying each value of x with the the probability of it occurring, “weighted average”

Variance

> https://www.youtube.com/watch?v=Vyk8HQOckIE

- The definition of “classical” variance from the previous section is the one you might be used to (at least I am) but the one Joel uses relies on expected values and I think its especially for discrete random variables, here it is:

$$\sigma^2 = E[(X-\mu)^2] = \sum (x-\mu)^2 \cdot p(x)$$

- Can also be rewritten as (this formulation is useful because its usually easier to calculate the variance this way)

$$\sigma^2 = E(X^2) - \mu^2$$

- He shows how you can derive this in the slides and video

> ❗ Its super helpful to watch this video https://www.youtube.com/watch?v=Vyk8HQOckIE and see each of the above values calculated once if they are confusing or you are not sure how they work

Variance Sum Law

> https://www.youtube.com/watch?v=GH7dwoXSD0s

- What is the variance of the sum of two numbers?
- There is a long derivation in the video and slides

$$\mathrm{Var}[X+Y] = \sigma_X^2 + \sigma_Y^2 + 2(E[XY] - E[X]E[Y])$$

- Where $\sigma^2$is the variance, $E[X]$ is the expected value
- Using covariance:

$$\begin{aligned}\mathrm{Var}[X + Y] &= \sigma_X^2 + \sigma_Y^2 + 2\mathrm{Cov}(X,Y)\\ &= \sigma_X^2 + \sigma_Y^2 + 2\rho_{X,Y}\sigma_X\sigma_Y\end{aligned}$$

$$\rho_{X,Y} = \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

- If we wanted $Var[X - Y]$ instead, we only flip where the red sign is
- If X and Y are independent the $2Cov(X, Y)$ term goes away

Variance of an Average

$Var[\bar{X}] = Var[\frac{X_1+X_2+...+X_n}{n}] = \frac{\sigma^2}{n}$

- Derivation in slides and clips but the core idea is that the variances of all the samples are the same because they are sampled from the same distribution, and the $2Cov(X, Y)$ terms goes away in the derivation because they are independent
- If we take the square root of that result, we get the standard deviation $\frac{\sigma}{\sqrt{n}}$

Central Limit Theorem

> https://www.youtube.com/watch?v=YAlJCEDH2uY

- "Means are normally distributed" even if the distribution they are sampled from is not
- Advantage/use case: we can interpret results/make experiments without caring too much about what the underlying distribution is
- A few caveats to the CLT
  - n the number of samples needs to be sufficiently large, whats sufficient? ask god

    - The more “not normal” the original distribution is, the more samples you need to arrive at our normal distribution of the means, he showed this with some visual examples in the slides
    - Conditions
        - Data must be sampled randomly
        - Data must be independent
        - Sample sizes should be no more than 10% of the population
        - Sample size sufficiently large

Monte Carlo Integration

- Something about needles ... but I dont think its important, its just explaining what Monte Carlo simulations are, and its basically when you use randomness in an experiment in a smart way, here is a nice video https://www.youtube.com/watch?v=7ESK5SaP-bc
- There is a derivation in the slides and video
- Its a “brute force” way to solve integrations instead of actually computing them analytically by hand (some can not be computed analytically by hand)

$$\int_a^b f(x)dx = (b-a)\frac{1}{N}\sum_{i=0}^{N-1} f(X_i).$$

- for n randomly sampled variables from normal distribution