## Conditional Expectation, Introductory Inference

Before you leave lab, make sure you click here so that you're marked as having attended this week's section. The CA leading your discussion section can enter the password needed once you've submitted.

## 1 Warmups

### 1.1 Why Multiple Random Variables?

What is a probabilistic model with multiple random variables? What does the term inference mean? What do you call the probability of an assignment to all variables in a probabilistic model? Why is that useful? Why can it be hard to represent?

### 1.2 Joint Random Variables Statistics

True or False? The symbol Cov is covariance, the symbol $\wedge$ is logical-and, the symbol $\rho$ is Pearson correlation, the symbol $\Longrightarrow$ is logical implication, and $X \perp Y$ is just a fancy way to say that $X$ and $Y$ are independent.

A statement like " $A \sim \operatorname{Bin}(10,0.5) \wedge B \sim \operatorname{Bin}(10,0.5) \wedge A \perp B \Longrightarrow A+B \sim \operatorname{Bin}(20,0.5) "$ reads "If $A$ and $B$ are both distributed as Binomials with the same parameters, then $A+B$ is a Binomial as well with the same $p$ parameter and an $n$ parameter that's the sum of the those for $A$ and $B$."

$$
\begin{array}{c|c}
X \perp Y \Longrightarrow \operatorname{Cov}(X, Y)=0 & \operatorname{Var}(X+X)=2 \operatorname{Var}(X) \\
\hline \operatorname{Cov}(X, Y)=0 \Longrightarrow \mathrm{X} \perp Y & X \sim \mathcal{N}(0,1) \wedge Y \sim \mathcal{N}(0,1) \Longrightarrow \rho(X, Y)=1 \\
\hline Y=X^{2} \Longrightarrow \rho(X, Y)=1 & Y=3 X \Longrightarrow \rho(X, Y)=3
\end{array}
$$

### 1.3 Random Number of Random Variables

Let $N$ be a non-negative integer-valued random variable-that is, a random variable that takes on values in $\{0,1,2, \ldots\}$. Let $X_{1}, X_{2}, X_{3}, \ldots$ be an infinite sequence of independent and identically distributed random variables (independent of $N$ ), each with mean $\mu$, and $X=\sum_{i=1}^{N} X_{i}$ be the sum of the first $N$ of them.

Before doing any work, what do you think $E[X]$ will turn out to be? Then show it mathematically to see if your intuition is correct.

## 2 Problems

### 2.1 CS106A Is Popular

CS106A is Stanford's introductory programming course and largely considered the primary gateway to our undergraduate major. Assume next quarter's offering of CS106A is exactly 600 people, that each of the four undergraduate classes is compromised of 1750 students, and that next quarter's CS106A roster is just some random sample of the 7000 undergraduates. Let $A, B, C$, and $D$ count the number of freshman, sophomores, juniors, and seniors in the class of 600.

- Present the joint probability mass function of $A, B, C, D$ ? Restated, present an expression for $P(A=a, B=b, C=c, D=d)$.
- Does your PMF from part a) describe a multinomial random variable? Intuitively justify your answer.
- What is the conditional probability distribution of $A$ given that $B+C=300$ ? Restated, what is $P(A=a \mid B+C=300)$ ?
- Do you expect $\operatorname{Cov}(A, B)$ to be positive, zero, or negative? Justify your answer.


### 2.2 Managing Screen Time

Push notifications light up our phones at a rate that's guided by a Poisson process with an constant average rate of 5 notifications per hour at all hours, night and day. In an effort to maximize productivity, you put your phone down and ignore it as much as possible. You do, however, periodically check it to clear all notifications. You check at 7 am when you wake up, noon when you grab lunch, 5 pm as you wrap up classes for the day, and then again at 10 pm before you go to sleep.

Let $\mathrm{W}, \mathrm{X}, \mathrm{Y}$, and Z be Poisson random variables that count the number of push notifications that have accumulated from 10 pm to 7 am , 7 am to noon, noon to 5 pm , and 5 pm to 10 pm , respectively. We assume that the number of push notifications that arrive within each interval are all mutually independent of all other intervals.

- Compute the joint PMF on W, X, Y, and Z.
- Compute the conditional joint PMF on $\mathrm{W}, \mathrm{X}, \mathrm{Y}$, and Z given that $\mathrm{W}+\mathrm{X}+\mathrm{Y}+\mathrm{Z}=150$.
- Compute the conditional PMF of $\mathrm{X}+\mathrm{Y}+\mathrm{Z}$-that's the number of notifications that arrive while you're awake-given that $\mathrm{W}+\mathrm{X}+\mathrm{Y}+\mathrm{Z}=150$.
- Compute $E[X+Y+Z \mid W+X+Y+Z=150]$ and $\operatorname{Var}(X+Y+Z \mid W+X+Y+Z=150)$.

|  | $\mathrm{A}=\mathbf{0}$ |  |  | $\mathrm{A}=\mathbf{1}$ |  |
| :---: | :---: | :---: | :--- | :--- | :--- |
|  | $\mathrm{B}=0$ | $\mathrm{~B}=1$ |  | $\mathrm{~B}=0$ | $\mathrm{~B}=1$ |
| $\mathrm{C}=0$ | 0.36 | 0.20 |  | 0.00 | 0.00 |
| $\mathrm{C}=1$ | 0.04 | 0.20 |  | 0.10 | 0.10 |

### 2.3 Understanding Bayes Nets

The joint probability table (above) for random variables $A, B$ and $C$ is equivalent to the Bayes network (below). Both give the probability of any combination of the random variables. In the Bayes network the probability of each random variable is provided given its causal parents.


- Use the Bayes network to explain why $P(A=0, B=1, C=1)=0.20$
- What is $P(A=1 \mid C=1)$ ?
- Is $A$ independent of $B$ ? Explain your answer.
- Is $A$ independent of $B$ given $C=1$ ? Explain your answer.

