Jerry Cain
CS 109

Section 5
May 4, 2022

Section 5 Solutions

Based on the work of many CS109 instructors and course staff members.

1 Problems

1.1 Joint Random Variables Statistics

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

A statement like "$X \sim \mathcal{N}(0,1) \wedge Y \sim \mathcal{N}(0,1) \implies \rho(X,Y) = 1$" reads "$X$ and Y both being distributed as standard normal distributions implies they are perfectly correlated".

$X \perp Y \implies Cov(X,Y) = 0$	$Var(X + X) = 2Var(X)$
$Cov(X,Y) = 0 \implies \text{X} \perp Y$	$X \sim \mathcal{N}(0,1) \wedge Y \sim \mathcal{N}(0,1) \implies \rho(X,Y) = 1$
$Y = X^2 \implies \rho(X,Y) = 1$	$Y = 3X \implies \rho(X,Y) = 3$

True or False?

True	False (... $= 4Var(X)$)
False (antecedent necessary, not sufficient)	False (don't know how independent X & Y are)
False ($Y = X \implies \dots$)	False (... = 1)

1.2 Hat-Check Again??

Recall the hat-check problem from an earlier discussion section: n people go to a party and drop off their hats to a hat-check person. When the party is over, a different hat-check person is on duty, and returns the n hats randomly back to each person. Let X be the random variable representing the number of people who get their own hat back. We showed last time that $E[X] = 1$ for any n.

What is $Var(X)$?

Similarly to last time, let $X_i \sim Bernoulli(p = 1/n)$ be the indicator variable for whether the i^{th} person gets their hat back, so that $X = \sum_{i=1}^{n} X_i$. Then,

$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

The first term is simply $np(1-p) = n\left(\frac{1}{n}\right)\left(1 - \frac{1}{n}\right) = \frac{n-1}{n}$ since each individual variance is $p(1-p)$.

To compute $Cov(X_i, X_j)$ for $i \neq j$, we note that $Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$. The random variable $X_i X_j$ is also Bernoulli, with parameter $1/n \cdot 1/(n-1)$ since both have to get their hat back. So

$$Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j] = \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2 = \frac{1}{n^2(n-1)}$$

Noting that there are $\binom{n}{2} = \dfrac{n(n-1)}{2}$ identical terms in the summation over $i < j$ and putting this all together gives

$$Var(X) = \frac{n-1}{n} + 2\binom{n}{2}\frac{1}{n^2(n-1)} = \frac{n-1}{n} + \frac{1}{n} = 1.$$

So $E[X] = Var(X) = 1$. What a coincidence!

1.3 Conditional Expectation

Let $X \sim Geo(p)$. Use the Law of Total Expectation to prove that $E[X] = 1/p$, by conditioning on whether the first flip is heads or tails.

$$E[X] = E[X|H]P(H) + E[X|T]P(T) = 1 \cdot p + (E[1+X])(1-p)$$

Solving yields $E[X] = 1/p$.

1.4 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 μ, 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.

$$E[X] = E\left[\sum_{i=1}^{N} X_i\right] = \sum_n E\left[\sum_{i=1}^{N} X_i \mid N = n\right] p_N(n) = \sum_n E\left[\sum_{i=1}^{n} X_i \mid N = n\right] p_N(n)$$

$$= \sum_n E\left[\sum_{i=1}^{n} X_i\right] p_N(n) = \sum_n n\mu p_N(n) = \mu \sum_n n p_N(n) = \mu E[N]$$

Alternatively,

$$E[X] = E[E[X|N]] = E[N\mu] = \mu E[N]$$

1.5 Binary Trees

Consider the following function for constructing binary trees:

```
struct Node {
    Node *left;
    Node *right;
};

Node *randomTree(float p) {
    if (randomBool(p)) {  // returns true with probability p
        Node *newNode = new Node;
        newNode->left = randomTree(p);
        newNode->right = randomTree(p);
        return newNode;
    } else {
        return nullptr;
    }
}
```

The `if` branch is taken with probability p (and the `else` branch with probability $1 - p$). A tree with no nodes is represented by `nullptr`; so a tree node with no left child has `nullptr` for the `left` field (and the same for the right child).

Let X be the number of nodes in a tree returned by `randomTree`. You can assume $0 < p < 0.5$. What is $E[X]$, in terms of p?

Let X_1 and X_2 be number of nodes the left and right calls to `randomTree`.
$E[X_1] = E[X_2] = E[X]$.

$$\begin{aligned}
E[X] &= p \cdot E[X \mid \texttt{if}] + (1-p)E[X \mid \texttt{else}] \\
&= p \cdot E[1 + X_1 + X_2] + (1-p) \cdot 0 \\
&= p \cdot (1 + E[X] + E[X]) \\
&= p + 2pE[X] \\
(1-2p)E[X] &= p \\
E[X] &= \frac{p}{1-2p}
\end{aligned}$$

2 Previous Exam Questions

2.1 Winter 2021: Quiz 2

You are cooking a big pot of vegetables and the instructions say to stir the pot". If you dont stir the pot some vegetables will be under-cooked and some will be overcooked.

How much heat a vegetable gets depends on if it is on the bottom of the pot or not. If it is on the bottom of the pot it will get 5 units of heat per minute. If it is not, it will get 1 unit of heat per minute.

Each time you stir, each vegetable, independently, has a 1/5 chance of ending up at the bottom of the pot. For each stir, the position of a vegetable is independent of previous position and of the position of other vegetables.[1]

For each of the questions below, in addition to providing an expression, please compute a numeric answer.

a. You stir the pot a single time and then let the vegetables cook for 10 minutes. What is the variance of heat on a vegetable?

Let X represent the units of heat a vegetable gets over the course of 10 minutes. Then,

$$E[X] = \frac{1}{5} \cdot (50) + \frac{4}{5} \cdot (10) = 18$$

and it follows that

$$Var(X) = \frac{1}{5} \cdot (50 - 18)^2 + \frac{4}{5} \cdot (10 - 18)^2 = 256$$

b. You stir the pot **once every minute** for 10 minutes. What is the variance of heat on a vegetable?

Let X_i represent the heat a vegetable receives at minute i where $1 \leq i \leq 10$. Then

$$E[X_i] = \frac{1}{5} \cdot (5) + \frac{4}{5} \cdot (1) = 1.8$$

and we find that

$$Var(X_i) = \frac{1}{5} \cdot (5 - 1.8)^2 + \frac{4}{5} \cdot (1 - 1.8)^2 = 2.56$$

To find the variance of the heat on a vegetable we must sum over the entire time period. Then, we want to find

[1]This independence assumption doesnt perfectly match the real world for example in the real world it wouldnt be possible for all vegetables to end up at the bottom of the pot. While the independence assumption is wrong, it has a very small impact on the final answer, and it makes for a more straightforward quiz.

$$Var\left(\sum_{i=1}^{10} X_i\right) = \sum_{i=1}^{10} Var(X_i) = \sum_{i=1}^{10} 2.56 = 25.6$$

c. Lets explore what this problem would look like with continuous numbers. In the continuous version, when you stir the pot, each vegetable has a distance from the bottom which is equally likely to be any real valued number between 0 and 1. The heat received by a vegetable, H, with distance d is: $H(d) = 1 - d^2$. You stir the pot a single time and then let the vegetables cook for 10 minutes. Show an expression that you would need to solve in order to calculate the expectation of heat. Your expression should include an integral. Then, solve the expression to get a numeric answer.

H is a rate measured in heat-units per minute. D is the distance measured in distance-units.

$$D \sim Uni(0, 1); p_D(d) = 1 \text{ for } 0 \leq d \leq 1, \text{ else } 0$$

$$H = g(D) = 1 - D^2$$

We want to compute $E[10H]$.

$$E[10H] = E[10g(D)] = 10E[g(D)]$$

Using the Law of the Unconscious Statistician:

$$\begin{aligned}
10E[g(D)] &= 10 \int_0^1 g(d) p_D(d) dd \\
&= 10 \int_0^1 (1 - d^2) \cdot (1) dd \\
&= 10 \int_0^1 (1 - d^2) dd \\
&= 10[d - \frac{d^3}{3}]_0^1 \\
&= 10\left(\frac{2}{3}\right) \\
&= \frac{20}{3}
\end{aligned}$$