CLisa Yan and Jerry Cain
CS 109

Section #4
October 13 - 14, 2020

Section 4 Solutions

Based on the work of many CS109 staffs

1 Warmups

1.1 Joint Distributions

1. Given a Normal RV $X \sim N(\mu, \sigma^2)$, how can we compute $P(X \leq x)$ from the standard Normal distribution Z with CDF ϕ?

2. What is a continuity correction and when should we use it?

3. If we have a joint PMF for discrete random variables $p_{X,Y}(x, y)$, how can we compute the marginal PMF $p_X(x)$?

> 1. First, we write $\phi((x - \mu)/\sigma)$. We then look up the value we've computed in the Standard Normal Table.
>
> 2. Continuity correction is used when a Normal distribution is used to approximate a Binomial. Since a Normal is continuous and Binomial is discrete, we have to use a continuity correction to discretize the Normal. The continuity correction makes it so that the normal variable is evaluated from + or - 0.5 increments from the desired k value.
>
> 3. The marginal distribution is $p_X(x) = \sum_y p_{X,Y}(x, y)$

1.2 Independent Random Variables

1. What distribution does the sum of two independent binomial RVs $X + Y$ have, where $X \sim Bin(n_1, p)$ and $Y \sim Bin(n_2, p)$? Include the parameter(s) in your answer. Why is this the case?

2. What distribution does the is of two independent Poisson RVs $X+Y$ have, where $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$? Include the parameter(s) in your answer.

3. If $Cov(X, Y) = 0$, are X and Y independent? Why or why not?

> 1. Binomial; $X + Y \sim Bin(n_1 + n_2, p)$
>
> 2. Poisson; $X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$

3. Not necessarily. Suppose there are three outcomes for X: let X take on values in $\{-1, 0, 1\}$ with equal probability $1/3$. Let $Y = X^2$. Then, $E[XY] = E[X^3] = E[X] = 0$ (since $X^3 = X$) and $E[X] = 0$, so $Cov(X, Y) = E[XY] - E[X]E[Y] = 0 - 0 = 0$ but X and Y are dependent since $P(Y = 1) = 2/3 \neq 1 = P(Y = 1|X = 1)$.

1.3 Joint Random Variables Statistics

1. **True or False?** The symbol Cov is covariance, and the symbol ρ is Pearson correlation.

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

True or False?

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

2 Problems

2.1 Approximating Normal

Your website has 100 users and each day each user independently has a 20% chance of logging into your website. Use a normal approximation to estimate the probability that more than 21 users log in.

The number of users that log in B is binomial: $B \sim \text{Bin}(n = 100, p = 0.2)$. It can be approximated with a normal that matches the mean and variance. Let C be the normal that approximates B. We have $E[B] = np = 20$ and $Var(B) = np(1-p) = 16$, so $C \sim N(\mu = 20, \sigma^2 = 16)$. Note that because we are approximating a discrete value with a continuous random variable, we need to use the continuity correction:

$$\begin{aligned} P(B > 21) &\approx P(C > 21.5) \\ &= P\Big(\frac{C-20}{\sqrt{16}} > \frac{21.5-20}{\sqrt{16}}\Big) \\ &= P(Z > 0.375) \\ &= 1 - P(Z < 0.375) \\ &= 1 - \phi(0.375) = 1 - 0.6462 = 0.3538 \end{aligned}$$

2.2 Are we due for an earthquake?

After the class where we talked about the probability of Earthquakes at Stanford, a student asked a question: "Doesn't the probability of an earthquake happening change based on the fact that we haven't had one for a while?" Let's explore! Recall the USGS rate of earthquakes of magnitude 8+ is λ =0.002 earthquakes per year.

a. What is the probability of no 8+ earthquakes in four years after the 1908 earthquake (recall that earthquakes are exponentially distributed)?

Let X be the time until an earthquake. $X \sim \text{Exp}(\lambda = 0.002)$.

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) \\ &= 1 - F_X(4) \\ &= 1 - [1 - e^{-0.002 \cdot 4}] \\ &= e^{-0.008} \approx 0.992 \end{aligned}$$

b. What is the probability of no 8+ earthquakes in the 113 years between the 1908 earthquake and four years from now?

$$
\begin{aligned}
P(X \geq 113) &= 1 - P(X < 113) \\
&= 1 - F_X(113) \\
&= 1 - [1 - e^{-0.002 \cdot 113}] \\
&= e^{-0.226} \approx 0.798
\end{aligned}
$$

c. What is the probability of no 8+ earthquakes in the 113 years between the 1908 earthquake and four years from now *given* that there have been no earthquakes in the last 109 years?

$$
\begin{aligned}
P(X > 113 | X > 109) &= \frac{P(X > 113, X > 109)}{P(X > 109)} \\
&= \frac{P(X > 113)}{P(X > 109)} = \frac{1 - F_X(113)}{1 - F_X(109)} \\
&= \frac{e^{-0.002 \cdot 113}}{e^{-0.002 \cdot 109}} = e^{-0.008} \approx 0.992
\end{aligned}
$$

d. Did you notice anything interesting? Would this work for any value of λ?

It turns out that exponentials are what we call a “memoryless distribution." If X is an exponential random variable, it holds that $P(X > s + t | X > t) = P(X > s)$.