Math 20: Probability

Midterm 2

Instructor: Xingru Chen

https://github.com/fudab/Math-20

July 20, 2020

Problem 1: True or False

10 = 2 + 2 + 2 + 2 + 2 pts

(a) True For any two numerically-valued random variable X and Y, if X and Y are independent, X^2 and Y^2 are also independent.

(b) False Let X be a random variable which can take on values $\{0, 1, 2, 3, \ldots\}$ (all non-negative integers). X can be uniformly distributed.

Hint: The sum of countable many zeros is still zero.

(c) False For a numerically-valued random variable X, if $E(X^2) = E^2(X)$, X can take two different values.

(d) False For any two numerically-valued random variable X and Y, if $\mathrm{COV}(X, Y) = 0$, X and Y are independent.

(e) False For any numerically-valued random variable X, its expected value and variance cannot be negative.

Problem 2: Computation

13 = 3 + 5 + 5 pts

Let X and Y be two **independent** random variables.

(a) Assume that X is the outcome of a single Bernoulli trial where

$$m(x) = \begin{cases} p, & X = 1 \\ 1-p, & X = 0 \end{cases}.$$

Find $E(X^{2020})$.

3 pts

$$E(X^{2020}) = 1^{2020} \times p + 0^{2020} \times (1-p) = p.$$

(b) Assume that X is Poisson distributed with parameter λ. Find $E(X^3)$.

5 pts

$$\begin{aligned}
E(X^3) &= \sum_{k=0}^{+\infty} k^3 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{+\infty} k^2 \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda \sum_{k=1}^{+\infty} k^2 \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{k=1}^{+\infty} (k^2-1) \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} + \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{k=2}^{+\infty} (k^2-1) \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} + \lambda \sum_{l=0}^{+\infty} \frac{\lambda^l}{l!} e^{-\lambda} \\
&= \lambda \sum_{k=2}^{+\infty} (k+1) \frac{\lambda^{k-1}}{(k-2)!} e^{-\lambda} + \lambda = \lambda^2 \sum_{l=0}^{+\infty} (l+3) \frac{\lambda^l}{l!} e^{-\lambda} + \lambda \\
&= \lambda^2 \sum_{l=0}^{+\infty} l \frac{\lambda^l}{l!} e^{-\lambda} + 3\lambda^2 \sum_{l=0}^{+\infty} \frac{\lambda^l}{l!} e^{-\lambda} + \lambda \\
&= \lambda^3 + 3\lambda^2 + \lambda.
\end{aligned}$$

(c) Assume that X and Y are exponentially distributed with parameter λ_1 and λ_2. Find $E((X+Y)^2)$.

5 pts

$$\begin{aligned}
E((X+Y)^2) &= E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2) \\
&= E(X^2) + 2E(X)E(Y) + E(Y^2) \\
&= \frac{2}{\lambda_1^2} + \frac{2}{\lambda_1 \lambda_2} + \frac{2}{\lambda_2^2}.
\end{aligned}$$

Problem 3: Proof

17 = 2 + 5 + 5 + 5 pts

Let X be the outcome of a chance experiment with $E(X) = \mu$ and $V(X) = \sigma^2$. When μ and σ are unknown, the statistician often estimates them by repeating the experiment n times with outcome $x_1, x_2, \cdots, x_n$, estimating μ by the sample mean

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

and σ^2 by the sample variance

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

Show that, for the sample mean $\bar{x}$ and sample variance s^2,

(a) $E(\bar{x}) = \mu$.

2 pts

$$E(\bar{x}) = E(\frac{1}{n}\sum_{i=1}^{n} x_i) = \frac{1}{n}E(\sum_{i=1}^{n} x_i) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu.$$

(b) $E((\bar{x} - \mu)^2) = \frac{\sigma^2}{n}$.

5 pts

Given that $E(\bar{x}) = \mu$,

$$\begin{aligned} E((\bar{x} - \mu)^2) = V(\bar{x}) &= V(\frac{1}{n}\sum_{i=1}^{n} x_i) \\ &= \frac{1}{n^2}V(\sum_{i=1}^{n} x_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} V(x_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 \\ &= \frac{\sigma^2}{n}. \end{aligned}$$

(c) $E(s^2) = \frac{n-1}{n}\sigma^2$.

5 pts

$$
\begin{aligned}
E(s^2) &= E[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2] = \frac{1}{n}E[\sum_{i=1}^{n}(x_i-\bar{x})^2] = \frac{1}{n}E[\sum_{i=1}^{n}(x_i-\mu+\mu-\bar{x})^2] \\
&= \frac{1}{n}E[\sum_{i=1}^{n}((x_i-\mu)^2-2(x_i-\mu)(\bar{x}-\mu)+(\bar{x}-\mu)^2)] \\
&= \frac{1}{n}E[\sum_{i=1}^{n}(x_i-\mu)^2-2\sum_{i=1}^{n}(x_i-\mu)(\bar{x}-\mu)+\sum_{i=1}^{n}(\bar{x}-\mu)^2] \\
&= \frac{1}{n}E[\sum_{i=1}^{n}(x_i-\mu)^2-2\sum_{i=1}^{n}(x_i-\mu)(\bar{x}-\mu)+\sum_{i=1}^{n}(\bar{x}-\mu)^2] \\
&= \frac{1}{n}E[\sum_{i=1}^{n}(x_i-\mu)^2-2n(\bar{x}-\mu)^2+n(\bar{x}-\mu)^2] = \frac{1}{n}E[\sum_{i=1}^{n}(x_i-\mu)^2-n(\bar{x}-\mu)^2] \\
&= \frac{1}{n}\sum_{i=1}^{n}E((X_i-\mu)^2)-E((\bar{x}-\mu)^2) = \frac{1}{n}\sum_{i=1}^{n}V(x_i)-V(\bar{x}) \\
&= \frac{n-1}{n}\sigma^2.
\end{aligned}
$$

The last part is independent of the above.

(d) Let X be a random variable with density function $f(x)$. Show, using elementary calculus, that the function

$$\phi(a) = E((X-a)^2)$$

takes its minimum value when $a = \mu(X)$, and in that case $\phi(a) = V(X)$.

5 pts

We know that $\phi(a) = E((X-a)^2) = \int_{-\infty}^{+\infty}(x-a)^2 f(x)dx$. Consider the derivative $\frac{d}{da}\phi(a)$:

$$\frac{d}{da}\phi(a) = \frac{d}{da}\int_{-\infty}^{+\infty}(x-a)^2 f(x)dx = -2\int_{-\infty}^{+\infty}(x-a)f(x)dx.$$

Furthermore the second-order derivative is $\frac{d^2}{da^2}\phi(a) = 2\int_{-\infty}^{+\infty} f(x)dx = 2 > 0$.

Therefore, $\phi(a)$ will take its minimum value when $\frac{d}{da}\phi(a) = 0$, that is,

$$\int_{-\infty}^{+\infty}(x-a)f(x)dx = \int_{-\infty}^{+\infty} xf(x)dx - a\int_{-\infty}^{+\infty} f(x)dx = \mu(X) - a = 0.$$

Now that we have $a = \mu(X)$, it follows that $\phi(a) = E((X-\mu(X))^2) = V(X)$.

Problem 4: Manipulation

10 = 5 + 5 pts

A number U is chosen at random in the interval $[0,1]$.

(a) Find the probability that $T = \sin(\pi U) < \frac{1}{2}$.

5 pts

Given that

$$\sin(\frac{\pi}{6}) = \sin(\frac{5\pi}{6}) = \frac{1}{2},$$

the probability that $T = \sin(\pi U) < \frac{1}{2}$ is equivalent to the probability that

$$0 \le \pi U < \frac{\pi}{6}, \qquad \text{or} \qquad \frac{5\pi}{6} < \pi U \le \pi.$$

Therefore,

$$P(T < \frac{1}{2}) = P((0 < U < \frac{1}{6}) \cup (\frac{5}{6} < U < 1)) = \frac{1}{3}.$$

(b) Let X be a random variable uniformly distributed over $[c,d]$. For what choice of a and b do we have $U = aX + b$?

5 pts

Consider the cumulative distribution function of $U = aX + b$. We have

$$F_U(u) = P(U \le u) = P(aX + b \le u) = P(X \le \frac{u-b}{a}).$$

Since U is uniformly distributed on $[0,1]$, $F_U(u) = u$, for $0 \le u \le 1$.

Also, as X is uniformly distributed over $[c,d]$, $F_X(x) = \frac{x-c}{d-c}$, for $c \le x \le d$.

Therefore, the original equation can be rewritten as

$$u = F_U(u) = P(X \le \frac{u-b}{a}) = \frac{\frac{u-b}{a} - c}{d-c}.$$

That is,

$$a(d-c)u = u - b - ac, \qquad \forall 0 \le u \le 1.$$

Comparing the coefficients, we get

$$a(d-c) = 1, \qquad b + ac = 0.$$

Simple calculation gives that

$$a = \frac{1}{d-c}, \qquad b = -\frac{c}{d-c}.$$

Problem 5: Educational attainment

10 = 5 + 5 pts

In August 2020, the Animal Bank released the Sharing Higher Education's Promise Report, highlighting the rising demand and supply of tertiary education. The Marmot Kingdom saw the fastest growth in its tertiary gross enrollment ratio (GER) during 2019 - 2020.

Back to 10 years ago, on the average, only 1 marmot in 1000 had a Bachelor degree.

(a) Find the probability that, in a county of 10,000 marmots, no one had a Bachelor degree.

5 pts

Let X be the number of marmots in the county having a Bachelor degree. We use a Poisson distribution with parameter

$$\lambda = np = 10000 \times \frac{1}{1000} = 10.$$

Then,

$$P(X = 0) = e^{-\lambda} = e^{-10}.$$

(b) How many marmots would have to be surveyed to give a probability greater than $\frac{1}{2}$ of finding at least one marmot with a Bachelor degree?

5 pts

Let n be the number of marmots taking the survey and X be the number of them having a Bachelor degree. This time we use a Poisson distribution with parameter

$$\lambda = np = \frac{n}{1000}.$$

And

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\lambda} = 1 - e^{-n/1000} > \frac{1}{2}.$$

Therefore,

$$n > 1000 \ln(2) \approx 693.1472.$$

Since n is an integer, $n_{\min} = 694$.

Problem 6: Cupid's Arrow

15 = 7 + 8 pts

It's said that Cupid carries two kinds of arrows, one with a sharp golden point, and the other with a blunt tip of lead. A person wounded by the golden arrow is filled with uncontrollable desire, but the one struck by the lead feels aversion and desires only to flee.

When Apollo taunts Cupid as the lesser archer, Cupid shoots him with the golden arrow, but strikes the object of his desire, the nymph Daphne, with the lead.

The duration of an arrow is the length of time that particular arrow is effective. The duration of the golden arrow has an exponential density with average time 200 years. And that of the lead arrow also has an exponential density but with average time 500 years.

golden arrow

lead arrow

Apollo wonders what is the probability that

(a) Daphne will break free from the enchantment first?

7 pts

Let X be the duration of the golden arrow and Y be that of the lead arrow. We have

$$f(x) = \frac{1}{200}e^{-x/200}, \qquad g(y) = \frac{1}{500}e^{-x/500}.$$

What we are looking for is the probability that

$$\begin{aligned}
P(X > Y) &= \int_{x=0}^{+\infty}\int_{y=0}^{x} f(x)g(y)dydx = \int_{x=0}^{+\infty} f(x)G(x)dx \\
&= \int_{x=0}^{+\infty} \frac{1}{200}e^{-x/200}(1-e^{-x/500})dx \\
&= \int_{x=0}^{+\infty} \frac{1}{200}e^{-x/200}dx - \int_{x=0}^{+\infty} \frac{1}{200}e^{-7x/1000}dx \\
&= 1 - \frac{5}{7} = \frac{2}{7}.
\end{aligned}$$

(b) himself will break free from the enchantment first and Daphne has to wait another 300 years or more before her arrow wears off?

8 pts

What we are now looking for is the probability that

$$
\begin{aligned}
P(Y \geq X + 300) &= \int_{x=0}^{+\infty} \int_{y=x+300}^{+\infty} f(x)g(y)dydx \\
&= \int_{x=0}^{+\infty} f(x)[1 - G(x+300)]dx \\
&= \int_{x=0}^{+\infty} \frac{1}{200} e^{-x/200} e^{-(x+300)/500} dx \\
&= \int_{x=0}^{+\infty} \frac{1}{200} e^{-3/5} e^{-7x/1000} dx \\
&= \frac{5}{7} e^{-3/5}.
\end{aligned}
$$

Problem 7: Man with No Name: A fistful of Nuts

15 = 7 + 8 pts

An unnamed stranger arrives at the little town of San Pecan on the Marmot-Squirrel border. He stays at Mr. Eyes closed's inn and befriends the aging innkeeper.

the Stranger

Mr. Eyes Closed

After settling down, the young and mysterious stranger opens a casino on the lobby.

A game is played as follows: a participant choose a random number X uniformly from $[0,1]$. Then a sequence $Y_1, Y_2, \cdots$ of random numbers is chosen independently and uniformly from $[0,1]$. The game ends the first time that $Y_i > X$ and the participant is paid $(i-1)$ pecans.

(a) Let $P(N = n \mid X = x)$ be the probability that the payout will be $N = n$ pecans given $X = x$. Which amount would the participant be most likely to get given $X = x$?

Hint from Mr. Eyes Closed: Which amount would the participant **expect** to get given $X = x$?

7 pts

We know that

$$P(N = n \mid X = x) = x^n(1-x),$$

which resembles a geometric distribution with $p = 1-x$.

Therefore,

$$E(N \mid X = x) = \sum_{n=0}^{+\infty} nx^n(1-x) = x\sum_{n=0}^{+\infty} nx^{n-1}(1-x) = \frac{x}{1-x}.$$

(b) What is a fair entrance fee for this game?

8 pts

Notice that the density function of X is $f(x) = 1$. Hence,

$$\begin{aligned}
E(N) &= \int_0^1 E(N|X = x) f(x) dx \\
&= \int_0^1 \frac{x}{1-x} dx \\
&= \int_0^1 (\frac{1}{1-x} - 1) dx \\
&= -\ln(1-x)|_0^1 \\
&= +\infty.
\end{aligned}$$

Therefore, a participant is willing to pay any amount of pecans for this game.

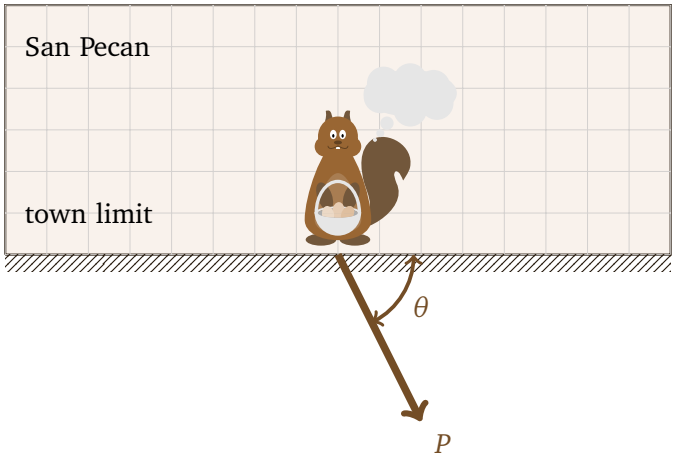
Problem 8: Man with No Name: Out of San Pecan

10 pts

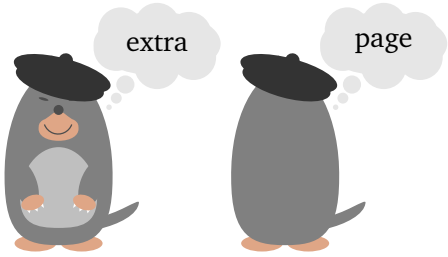
Right across from the inn live the town sheriff Butter Beer, his wife Puff, and their son Nutella. As the only child in the family, Nutella was always under the umbrella of his parents and has never traveled far. However, after hearing about the college-going culture in the Marmot Kingdom, he decides to leave home and make his own way in the faraway places.

Butter Beer Nutella Puff

Now Nutella is standing on the town limit. He chooses, at random, a direction which will lead him to the world, and walks 2 miles in that direction. Let P denote his position. What is the expected distance from P to the town limit?



Hint from the stranger: θ is a random variable.



10 pts

Consider the random variable θ uniformly distributed on $[0, \frac{\pi}{2}]$. The density function is $f(\theta) = \frac{2}{\pi}$.

We know that the distance from P to the town limit is

$$\phi(\theta) = 2\sin(\theta).$$

Therefore, the expected distance is

$$\begin{aligned} E(\phi(\theta)) &= \int_0^{\frac{\pi}{2}} \phi(\theta) f(\theta) d\theta \\ &= \int_0^{\frac{\pi}{2}} 2\sin(\theta)\frac{2}{\pi} d\theta \\ &= -\frac{4}{\pi}\cos(\theta)|_0^{\frac{\pi}{2}} \\ &= \frac{4}{\pi}. \end{aligned}$$