Math 372: Homework 06

Due Monday, October 12

Problem 1 (Variance shortcut formula)**.** Let X be a random variable, and let $\mu = \mathbb{E}[X]$. By definition of variance, we know that $\mathrm{Var}(X) = \mathbb{E}\left[(X-\mu)^2\right]$. Show that $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \mu^2$.

Problem 2. Let K be the number of Prussian cavalry officers killed by horsekick in a random year. Because of careful data collection by the Prussian military[1], we know that K has the following distribution:

$$p(0) = .51, \quad p(1) = .33, \quad p(2) = .11, \quad p(3) = .04, \quad p(4) = .01, \quad \text{and} \quad \mathbb{P}[K \geq 5] = 0.$$

(a) What is the expectation of the random variable K?

(b) What is the variance of K?

(c) Let $\lambda = \mathbb{E}[K]$, rounded to 1 decimal place, which you computed in part (a). Using this value, let Y be a Poisson random variable with rate parameter λ, and let $q(y) := \mathbb{P}[Y = y]$ be the pmf of Y. Compute $q(0), q(1), q(2), q(3), q(4)$, and $\mathbb{P}[Y \geq 5]$, and compare the probability mass functions p and q.

Problem 3. Given an example with dice that show that, in general,

$$\mathbb{P}(B) \neq \mathbb{P}(B|A) + \mathbb{P}(B|A^c).$$

Problem 4. Compute $\mathbb{E}[X]$, $\mathbb{E}\left[X^2\right]$ and $\mathrm{Var}(X)$ for the following random variables:

(a) $X =$ (value of one roll of a 6-sided dice).

(b) A fair coin toss. $Y = \begin{cases} 1 & : \text{ with probability } \frac{1}{2} \\ 0 & : \text{ with probability } \frac{1}{2} \end{cases}$

(c) A biased coin toss. $Z = \begin{cases} 1 & : \text{ with probability } p \\ 0 & : \text{ with probability } 1-p \end{cases}$

Problem 5. Alice rolls six dice and wins if she rolls at least one six. Bob rolls twelve dice and wins if he rolls at least two sixes. Who has the greater probability of winning? *[Hint: calculate the probabilities to lose.]*

Problem 6. A certain brand of upright freezer is available in three different rated capacities: 16, 18, and 20 cubic feet. Let X be the rated capacity of a freezer of this brand sold at a certain store. Suppose that X has probability mass function

x	16	18	20
$p(x)$	.2	.5	.3

(a) Compute $\mathbb{E}[X], \mathbb{E}\left[X^2\right], \mathrm{Var}(X)$.

(b) If the price of a freezer having capacity X is $70X - 650$, what is the expected price paid by the next customer to buy a freezer?

(c) What is the variance of the price paid by the next customer?

(d) Suppose that although the rated capacity of a freezer is X, the actual capacity is $h(X) = X - .008X^2$. What is the expected actual capacity of the freezer purchased by the next customer.

Problem 7 (Exploding Dice)**.** Compute the expectation of an exploding 6-sided dice. ***Note:*** *An 'exploding' dice roll is a mechanic used in various tabletop games whereby the player rolls a dice, and then if they rolled the maximum number on the die, they get to roll it again and add the new roll to the old one. This can happen more than once.*

[1] This problem actually comes from real data collected by the Prussian army, and was one of the first known applications of the Poisson distribution, by Russian mathematician L. Bortkewitsch in 1898; see `https://www.jstor.org/stable/2348169?seq=3`

Problem 8 (Geometric random variable)**.** A basketball player is shooting hoops. Suppose that for each shot, the probability of successfully making the shot is $p \in (0,1)$, (where p depends on the skill of the player), and the probability of missing the shot is $1-p$. Let

$$X = (\#\text{ of throws until they make a basket})$$

(a) Explain why the pmf of X is

$$\mathbb{P}[X=k] = \begin{cases} (1-p)^{k-1}p & : \quad k = 1,2,\dots \\ 0 & : \quad \text{otherwise} \end{cases}$$

(A random variable with this pmf is called a ***geometric random variable*** with success probability p, and we write $X \sim \text{Geom}(p)$. The parameter p is called the ***success probability***.)

(b) Compute $\mathbb{P}[X > t]$ for $t \in \mathbb{R}$ (hint: consider the cases of $t > 0$ and $t < 0$ separately), and use your answer to find the cdf of X

$$F(t) := \mathbb{P}[X \le t], \qquad t \in \mathbb{R}.$$

(c) Plot the cdf F of X for $p = 1/2$. (Note: it should be a nondecreasing step function with $F(-\infty) = 0$ and $F(+\infty) = 1$.)

(d) Suppose that X is a geometric random variable with success parameter $p \in (0,1)$. Show that

$$\mathbb{P}[X = m+k \mid X > m] = \mathbb{P}[X = k] \quad \text{for integers } m > 0 \text{ and } k > 1.$$

Explain why this is called the *memoryless property* of the geometric distribution.

Problem 9. The human genome consists of 3.2 billion nucleotides (`A`, `G`, `C`, `T`). For this problem we can think of the human genome as a sequence of nucleotides

`CAGGAAGCATTTATATAGAGGTAATAATTAAATTTAACTTTTTATTATTCTCTATGCCCAAGAGAATGTG...`

consisting of 3,200,000,000 nucleotide letters. The DNA mutation rate in humans is approximately $2\times10^{-8} = 0.00000002$ mutations per nucleotide per generation. Assume that nucleotides mutate *independently*, that is, whether one nucleotide undergoes mutates is independent from any other nucleotides.

Let X be the number of mutations in the genome of a randomly selected newborn baby.

(a) Explain why X is a binomial random variable. Identify the value of the two parameters n (number of trials) and p (success probability).

(b) The average number of mutations in a newborn baby's genome is $\mu := \mathbb{E}[X]$. Find μ.

(c) For statistical reasons that we will get to later in the class,[2] over 95% of babies will have a number of mutations in the interval

$$(\mu - \sigma, \mu + \sigma)$$

where $\sigma = \sqrt{\text{Var}(X)}$. Find σ and (using your answer to (b)) find the interval $(\mu-\sigma, \mu+\sigma)$.

(d) Attempt to compute some probabilities, such as $\mathbb{P}[X = 1]$ and $\mathbb{P}[X = 60]$, and $\mathbb{P}[X \le 100]$. I do not expect you to be successful with this. What makes this task difficult?

(e) Let Y be a Poisson distribution with mean μ. Explain why Y can be used to approximate X, and compute the following probabilities:

(i) $\mathbb{P}[Y = 0]$
(ii) $\mathbb{P}[Y = 1]$
(iii) $\mathbb{P}[Y = 5]$
(iv) $\mathbb{P}[Y \ge 3]$

Problem 10. Let Y be a Poisson random variable with parameter $\lambda > 0$. Find $\mathbb{E}\left[\frac{1}{Y+1}\right]$.

[2] The Central Limit Theorem