Name: ______________________ PU ID: ______________

ECE 302: Probabilistic Methods in Electrical and Computer Engineering
Fall 2021
Instructor: Prof. A. R. Reibman

PURDUE UNIVERSITY

Homework 5

Fall 2021
(Due Thursday October 7, 11:59pm)

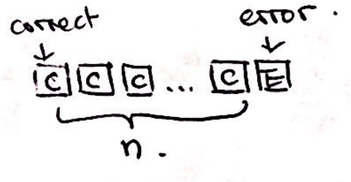
Homework is due on **Thursday October 7 at 11:59pm** on Gradescope. No late homework will be accepted, and **no homework will be accepted without a statement**. Include a brief description of all sources of information you used (including other people), not counting the text, handouts, or material posted on the web page, **or** state "I did not receive help on this homework". You do not need to reference any material presented in class or on the course web-site, in the textbook, nor Prof. Reibman nor TA Haoyu Chen.

Statement:

Topics: Common PMFs (Chapter 3.5); Conditional PMF, PDF, CDF (Chapters 3.4 and 4.2.2);

Exercise 1. (From textbook, problem 3.52)
A sequence of characters is transmitted over a channel that introduces errors with probability $p = 0.01$.

(a) What is the pmf of N, the number of error-free characters between erroneous characters?



$$P(N=n) = 0.99^n \cdot 0.01 \quad , \text{ for } n = 0, 1, \ldots$$

$$= (1-0.01)^n \cdot 0.01 .$$

Exercise 1, continued.

(b) What is $E(N)$?

(c) Suppose we want to be 99% sure that at least 1000 characters are received correctly before a bad one occurs. What is the appropriate value of p to achieve this?

b) note that N can be viewed as geometric RV for n consecutive success.

$$\therefore E(N) = \sum_{n=0}^{\infty} n \cdot P(N=n) = \frac{1-p}{p} = \frac{1-0.01}{0.01} = 99.$$

(geometric series sol:

$$= p(1-p) \cdot \sum_{n=1}^{\infty} n \cdot (1-p)^{n-1} = p(1-p) \frac{1}{(1-(1-p))^2} = \frac{1-p}{1}.$$

c). we want $P(N \geq 1000) = 0.99$.

$$\sum_{n=1000}^{\infty} P(N=n) = 0.99$$

$$\boxed{\sum_{n=0}^{k} a^n = \frac{1-a^{k+1}}{1-a}}$$

$$\sum_{n=1000}^{\infty} p(1-p)^n = 0.99.$$

$$\sum_{n=0}^{\infty} p \cdot (1-p)^n - \sum_{n=0}^{999} p(1-p)^n = 0.99$$

$$1 - p \cdot \frac{1-(1-p)^{1000}}{1-(1-p)} = 0.99.$$

$$1-p = 0.99^{\frac{1}{1000}}$$

$$p = 1 - 0.99^{\frac{1}{1000}} \approx 1 \times 10^{-5}$$

Exercise 2. (FROM EXAM 2, FALL 2015)
Five cars start out on a cross-country race. The probability that a car breaks down and drops out of the race is 0.2. Cars break down independently of each other.

(a) What is the probability that exactly two cars finish the race?

(b) What is the probability that at most two cars finish the race?

(c) What is the probability that at least three cars finish the race?

Binomial RV, p(finish the race) = p = 0.8.

a) $P(N=2) = \binom{5}{2} \cdot 0.2^3 \cdot 0.8^2 = \frac{5!}{2!\,3!} \cdot 0.2^3 \cdot 0.8^2$

$= 0.0512$.

b) $P(N \le 2) = \sum_{n=0}^{2} \binom{5}{n} p^n \cdot (1-p)^{5-n}$

$= \binom{5}{0} \cdot 0.2^5 + \binom{5}{1} \cdot 0.8 \cdot 0.2^4 + \binom{5}{2} \cdot 0.8^2 \cdot 0.2^3$

≈ 0.058.

c) $P(N \ge 3) = 1 - P(N \le 2) = 0.942$.

alternatively

$P(N \ge 3) = \sum_{n=3}^{5} \binom{5}{n} \cdot p^n \cdot (1-p)^{5-n}$

$= \binom{5}{3} p^3 (1-p)^2 + \binom{5}{4} p^4 (1-p) + \binom{5}{5} p^5$

Exercise 3. (FROM TEXTBOOK, PROBLEM 3.66)
A data center has 10,000 disk drives. Suppose that a disk drive fails in a given day with probability 10^{-3}.

(a) Find the probability that there are no failures in a given day.

(b) Find the probability that there are fewer than 10 failures in 2 days.

(c) Find the number of spare disk drives that should be available so that all failures in a day can be replaced with probability 99%.

a.) Poisson RV with $\alpha = 10000 \cdot 10^{-3} = 10$.

$$P(X=0) = \frac{\alpha^0 \cdot e^{-\alpha}}{0!} = e^{-10}.$$

b.) let Y be number of failures in 2 days.

$\therefore$ Poisson RV with $\alpha' = 20$.

$$P(Y<10) = \sum_{k=0}^{9} \frac{\alpha'^k e^{-\alpha'}}{k!} = 0.005.$$

c.) Find a number N such that

$P(X \leq N) = 0.99$.

$$\sum_{k=0}^{N} \frac{\alpha^k \cdot e^{-10}}{k!} = 0.99$$

Solve for $N = 18$.

Exercise 4. (From textbook, problem 3.53(a) and 3.54)
Let M be a geometric random variable with $S_M = \{1, 2, 3, \ldots\}$.

(a) Find $P(M = k | M \leq n)$.

(b) Show that M satisfies the memoryless property:

$$P(M \geq k + j | M \geq j + 1) = P(M \geq k) \quad \text{for all } j, k > 1$$

a.) $P(M=k | M \leq n) = \dfrac{P(\{M=k\} \cap \{k \leq n\})}{P(M \leq n)}$

$$= \begin{cases} \dfrac{P(M=k)}{P(M \leq n)}, & \text{for } k \leq n \\ 0, & \text{for } k > n. \end{cases}$$

$$= \begin{cases} \dfrac{(1-p)^{k-1} \cdot p}{1-(1-p)^n}, & \text{for } k \leq n \\ 0, & \text{for } k > n. \end{cases}$$

b.) $P(M \geq k+j | M \geq j+1) = \dfrac{P(\{M \geq k+j\} \cap \{M \geq j+1\})}{P(M \geq j+1)}$.

$\because k > 1$

$\therefore \{M \geq k+j\}$ ensures that $\{M \geq j+1\}$

$\therefore P(M \geq k+j | M \geq j+1) = \dfrac{P(M \geq k+j)}{P(M \geq j+1)} = \dfrac{1 - P(M \leq k+j-1)}{1 - P(M \leq j+1-1)}$.

$$= \frac{1-(1-(1-p)^{k+j-1})}{1-(1-(1-p)^{j})} = (1-p)^{k-1}$$

$$= 1 - P(M \leq k-1)$$

$$= P(M \geq k).$$

Exercise 5.
The goal of this exercise is to explore the question "How large is 'large n' and how small is 'small p'?" when the Poisson RV is used to approximate a Binomial RV, with $\alpha = np$.

Write a program in the language of your choice. Matlab has built-in functions **binopdf**, **binocdf**, **poisspdf**, and **poisscdf** that can be helpful, and I expect Python and Octave have similar functions. Also, they are quite straightforward to implement from scratch if you choose. Attach the plots and programs you generate to your homework. Be sure to include labels on the plots. Describe in words the conclusions that you draw from the experiment.

(a) Compute the minimum value of n for which the cumulative error between the PDFs $\sum_{k=0}^{n} |p_X(k) - p_Y(k)| < r$, where $r = 0.05, 0.01, 0.001$, and the RV X is a Binomial RV with parameters (n, p) and the RV Y is a Poisson RV with parameter $\alpha = np = 1$.

(b) Compute the minimum value of n for which the maximum error between the CDFs $\max_{k\in\{0,1,\ldots,n\}} |F_X(k) - F_Y(k)| < r$, where $r = 0.05, 0.01, 0.001$, and the RV X is a Binomial RV with parameters (n,p) and the RV Y is a Poisson RV with parameter $\alpha = np = 1$.

(c) Rerun, and report your results, for $\alpha = np = 2$ and $\alpha = np = 5$.

a.) find minimum n required to achieve desired threshold r.

c.)

	$\alpha=1$	$\alpha=2$	$\alpha=5$
$r<0.05$	$n=12$	$n=20$	$n=52$
Threshold $r<0.01$	$n=56$	$n=92$	$n=248$
$r<0.001$	$n=553$	$n=904$	$n=2452$.

- with the same rate α, larger n leads to more accurate estimation
- Smaller α tend to be easier to estimate.

b.) minimum n.

c.)

	$\alpha=1$	$\alpha=2$	$\alpha=5$
$r<0.05$	5	7	17
$r<0.01$	19	29	73
$r<0.001$	185	272	705.

- with same α, n, and p, the cumulative error between PDFs is larger than max error between CDFs.
i.e., it requires larger n to achieve same error threshold for cumulative error.

Conclusion: larger n leads to better estimation when using Poisson distribution to model a Binomial RV.

One should consider using Poisson RV for estimating Binomial RV when n is large enough (and p is small enough).