Midterm # 2.

Math 224
Introduction to Probability and Statistics

---

1. [binomial] We flip two identical fair coins simultaneously and record the configuration, which can be one of HH, HT, TH or TT. We do this trial 20 times in a row. Let X be the number of double heads HH we obtain at the end.

a) Identify the distribution of X, its mean value μ and its variance σ^2 by using the formulas.

b) What is the probability that we obtain three double heads HH?

Solution.

a) $p = .25$, $q = .75$, $n = 20$ X is the number of successes in a run of n bernoulli trials. The probability of getting a double head HH is 1/4. Then $X \sim BIN(20, .25)$.

b)

$$P(X = 3) =_{20} C_3 0.25^3 0.75^{17}$$

Finish the numerical computation.

2. [geometric, negative binomial] A basketball player can make a free throw with a constant probability $p = 1/4$. Let Y be the number of free throws needed in order to make his first shot. Let X be the number of free throws he must attempt in order to make <u>four</u> shots.

a) Determine the range (outcome space S), the p.m.f. of both Y and X and the name of their distributions.

b) Give the mean and variance of both X and Y.

c) Calculate $P(Y > 20)$ and $P(X = 8)$.

3. [negative binomial] Out of a shuffled regular deck of 52 cards, we draw one card at random. We record its face value and then put it back. Our goal is to successfully choose three aces. Let Y be the number of trials we need to perform in order to achieve our goal.

(a) What is the distribution of Y?

(b) What is the probability that $Y = 39$?

Solution. There are four aces in a deck of 52 cards; the odds of success from one trial (select one card) is $\frac{4}{52} = \frac{1}{13}$.

a) We have a negative binomial distribution $Y \sim NBIN(r = 3, p = \frac{1}{13})$.

b)

$$P(Y = 39) =_{39-1} C_{3-1} (\frac{1}{13})^3 (\frac{12}{13})^{39-3}$$

finish the computation.

4. [Poisson process] The number of telephone calls arriving at a customer center is modeled by a Poisson process $X(t)$ with rate 180 per hour. What is the probability that no more than two calls arrive during the first 80 seconds ?

Convert to seconds.

Solution. $\lambda \cdot 3,600 = 180$ (measured per second) then $\lambda = 0.05 = \frac{1}{20}$. Then, since the number of calls is a Poisson process with rate λ we know that the number

Y of calls in the time interval $[0, 80]$ (in seconds) is Poisson with parameter $\lambda \cdot 80 = 4$. Then

$$P(Y \leq 2) = \sum_{x=0}^{2} \frac{4^x}{x!} e^{-4} = 13e^{-4} = 0.238\,.$$

5. [Poisson approximation] A car insurance company with $n = 10,000$ policies issued has determined that the probability that a customer files a claim is $p = 0.0006$. Determine the probability distribution of the number of claims, which is denoted by X. Calculate $P(X = 7)$.

Solution. Each policy holder could make a claim with a small probability p and there are n policy holders. The number of claims X is exactly $X \sim Bin(n, p)$ which can be approximated by $X \sim Poisson(\lambda)$, where we take $\lambda = 6$, the average number of claims (that is, $\lambda = np = \mu_{binomial}$). Then

$$P(X = 7) = \frac{\lambda^7}{7!} e^{-\lambda} = \frac{6^7}{7!} e^{-6}$$

finish the numerical computation.

6. [Poisson process, exponential] Let $R(t)$ be a Poisson process with intensity $\lambda = 2$.

a) What is the distribution of X the number of arrivals (events, changes) occurring in the time interval $[0, 4]$? Give its name and exact formula, mean and variance *without doing the calculations.*

b) Give the distribution of W the waiting time between two consecutive arrivals. Justify informally based on the mean value. Write down the exact pdf. Determine $P(1 \leq W \leq 3)$.

7. [continuous distributions] Let X have the probability density function

$$f(x) = \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} .$$

Determine the median $m = \pi_{0.5}$, the third quartile $q_{0.75}$ and the 92nd percentile of X.

Solution. The c.d.f.

$$F(x) = \int_0^x f(z)dz = \int_0^x 2(1-z)dz = -(1-z)^2\Big|_0^x = 1 - (1-x)^2$$

Then the $100p$ percentile π_p is the solution to $F(\pi_p) = p$ that is

$$\pi_p = 1 - \sqrt{1-p}\,.$$

Apply this to
- $p = 0.5$ we get $m = \pi_{0.5} = 0.292$
- $p = 0.75$ we get $q_3 = \pi_{0.75} = 0.500$
- $p = 0.92$ we get $\pi_{0.92} = 0.717$.

8. [continuous distributions] Let X have the probability density function

$$f(x) = \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} .$$

Determine the mean μ and the variance σ^2.

9. [**continuous distributions**] Determine the constant c that makes $f(x) = cx^2$ on $[-1, 2]$ and zero otherwise a proper pdf.

a) Find the CDF $F(z)$, the median, first and second quartiles, and the formula for any p- quantile.

b) Determine the mean and variance of X, a random variable with distribution given by $f(x)$.

10. [**exponential**] Let $W \sim EXP(2)$.

– Prove that $P(W > 5 \,|\, W > 2) = P(W > 3)$.

– determine $P(W > 3)$.

– determine μ, σ^2 of W.

– Show that $F(z) = 1 - e^{-2z}$, $z > 0$.

11. [**normal**] The age X of a typical subscriber to a certain newspaper is a normal random variable with mean value 35.5 years old and standard deviation 4.8 years. What is the probability that X is between 30 and 40 years old?

Use the nearest values in the table.

Solution. We have $X \sim N(35.5, 4.8^2)$ then

$$P(30 < X < 40) = P\Big(\frac{30 - 35.5}{4.8} < Z < \frac{40 - 35.5}{4.8}\Big) =$$
$$\Phi(0.9375) - \Phi(-1,1458) \simeq \Phi(0.9375) + \Phi(1,1458) - 1 =$$
$$0.8264 + 0.8749 - 1 = 0.70\,.$$

12. [**normal**] Suppose that $X \sim N(3, 0.16)$. Find the following probabilities:

a) $P(X > 3)$

b) $P(2.8 < X < 3.1)$

c) Find the 98th percentile of X. Use Table V-a.

Solution.

$$\frac{X - 3}{\sqrt{.16}} = Z \sim N(0, 1)\,.$$

a)

$$P(X > 3) = P(\frac{X - 3}{.4} > \frac{3 - 3}{.4}) = 1 - \Phi(0) = .50$$

b)

$$P(2.8 < X < 3.1) = P(\frac{2.8 - 3}{.4} < \frac{X - 3}{.4} < \frac{3.1 - 3}{.4})$$
$$= \Phi(.25) - (1 - \Phi(.50)) = .3085$$

c)Let c be such that $P(X \leq c) = .98$, then

$$\frac{c - 3}{.4} = z_{0.02} \quad \Rightarrow \quad c = 3 + (.4)(2.055) = 3.822\,.$$