STAT201A: Introduction to Probability at an Advanced Level (Fall 2024)
UC Berkeley

Practice Problems for the Midterm Exam

Note: You are not expected to solve all these problems in just 80 minutes.

1. Determine whether each of the following claims is true or false. Provide reasons in each case.

   (a) It is often said that $\text{Bin}(n, p)$ is well-approximated by the $N(np, np(1-p))$ distribution. When $n = 3710$ and $p = 0.2$, this would mean that $\text{Bin}(3710, 0.2)$ is well-approximated by $N(742, 593.6)$. Therefore
   $$\frac{\mathbb{P}(\text{Bin}(3710, 0.2) \geq 941)}{\mathbb{P}\{N(742, 593.6) \geq 941\}}$$
   should be close to 1 (you might note here that $941/3710 \approx 0.254$).

   (b) Suppose X has the Negative Binomial distribution with parameters k and p (for example, X can be thought of as the distribution of the number of independent tosses of a coin with probability of heads p required to get the k^{th} head). Let $F_X(\cdot)$ denote the cdf of X. Then $F_X(X)$ has the uniform distribution on $(0, 1)$.

   (c) Suppose X has the geometric distribution with parameter $p(X$ can be thought of as the number of independent tosses of a coin with probability of heads p to get the first head). Then
   $$\mathbb{P}\{X > 3.5 + 1.5 \mid X > 1.5\} = \mathbb{P}\{X > 3.5\}$$

   (d) We can generate a random variable having any specified distribution by first generating a uniformly distributed random variable on $(0, 1)$ and then by applying an appropriate transformation to the uniform random variable.

2. Short questions.

   (a) There are 8 parents, 24 students and 3 teachers in a room. If a person is selected at random, what is the probability that it is a teacher or a student?

   (b) Find the probability to see 3 or less tails in 4 flips of a coin.

   (c) Suppose that A and B are independent, $\mathbb{P}(A) = 1/3$ and $\mathbb{P}(B) = 1/7$. Calculate $\mathbb{P}(A \cap B^c)$.

   (d) Suppose a box has 4 red marbles and 3 black ones. We select 2 marbles. What is the probability that second marble is red given that the first one is red?

   (e) Suppose the random variable X has possible values $\{1, 2, 3\}$ and probability mass function of the form $\mathbb{P}(X = k) = ck$. Find c. Find $\mathbb{E}[X]$. Find $\mathrm{Var}(X)$.

   (f) Let X be a random variable with exponential distribution with parameter 2. Find $\mathbb{P}(X > 14 \mid X > 4)$.

   (g) Russel has a biased coin for the which the probability of getting tails is an unknown p. He decide to flip the coin n and writes the total number of times X he gets tails. How large should n be in order to know with at least 0.95 certainty that the true p is within 0.1 of the estimate X/n? What if he wants 0.99 certainty?

   (h) Let X and Y be independent random variables with exponential distribution with parameter λ, find $\mathbb{P}(X > Y)$.

   (i) Let X be a random variable with m.g.f. $M_X(t) = e^{5t} - e^{3t}$. Find a formula for the moments of X.

   (j) Let X be a non-negative random variable with $\mathbb{E}[X] = 2$ and $\mathbb{E}[X^2] = 5$. Use Markov's inequality to find an upper bound for $\mathbb{P}(X > 10)$. Use Chebyshev's inequality to find an upper for $\mathbb{P}(X > 10)$.

3. Consider the urn setting that we discussed in lecture. We have an urn with R red balls and $N - R$ white balls. We draw balls in sequence from the urn without replacement.

   (a) Calculate $\mathbb{P}(F)$ where F denotes the proposition that the first red ball is drawn before the third white ball.

   (b) Calculate $\mathbb{P}(E)$ where E denotes the proposition that, when we draw n balls, our sample contains at least one red ball and at least two white balls.

4. Take random variables $X_1, X_2, X_3, \dots$ such that each of them has mean μ and variance 1.

   (a) Suppose that X_i are *negatively correlated*, i.e. $\mathrm{Cov}(X_i, X_j) < 0$ for all i, j. Set $S_n = X_1 + \cdots + X_n$. Show that (IMPORTANT: X_i are not independent!)

   $$\mathrm{Var}\left(\frac{S_n}{n}\right) \le \frac{1}{n}. \tag{1}$$

   (b) Assume instead that X_i are *positively correlated*, i.e. $\mathrm{Cov}(X_i, X_j) > 0$ for all i and j. Is (1) still true? Either give a proof or provide a counterexample.

5. (a) In Bernoulli (p) trials let V_n be the number of trials required to produce either n successes or n failures, whichever comes first. Find the distribution of V_n.

   (b) Suppose n balls are thrown independently at random into b boxes. Let X be the number of boxes left empty. Find expressions for $\mathrm{E}[X]$ and $\mathrm{Var}(X)$.

6. Suppose X and Y are independent random variables with X having the Exponential distribution with rate parameter λ and Y having the Standard Cauchy distribution. Let

$$U := \frac{Y\sqrt{X}}{\sqrt{1+Y^2}} \quad \text{and} \quad V := \frac{\sqrt{X}}{\sqrt{1+Y^2}}$$

   (a) Find the joint density of U and V.

   (b) Find the marginal densities of U and V.

   (c) Are U and V independent? Why or why not?