Review problems, Mathematical Statistics, Spring 2018

The following problems are intended to help you review the concepts you learned in Theory of Probability that you need to know for this course. You should try to write out your solutions in full, to be sure you understand them completely. This problem set was adapted from the 2016 version made by Miranda Holmes-Cefron.

A table of values of the cdf of the standard normal distribution is at the back.

1 Probability spaces

1. Give the definition of a sample space.
2. Suppose that the probability a man votes conservative is 0.45, the probability a woman votes conservative is 0.4, and the probability a woman votes conservative given her husband does is 0.6. What is the probability (a) both vote; (b) a man votes conservative given his wife does?
3. You are about to interview for NYU Law school. You know that 60% of interviewers are liberal, and 40% are conservative. You also know that 50% of liberals drink coffee but only 25% of conservatives do. Your interviewer makes both of you a cup of coffee. What is the probability she is a liberal?
(To answer this, use Bayes' Formula, and identify the events you are using.)
4. Three students acting independently each have probability 1/3 of solving a problem. What is the probability at least one of them will solve the problem?
5. Give the definition of a random variable.
6. Suppose we pick a letter uniformly at random from the work TENNESSEE. Which of the following is a random variable?
   (a) The letter that we pick.
   (b) The set $\{E, N, T, S\}$.
   (c) The probability that the letter is N.
   (d) The number of straight lines needed to write the letter.
   (e) The position of the letter in the alphabet (i.e. if we pick E, the position is 5, etc.)
7. A random variable Z has the following probability mass function: $p(-1) = 1/5$, $p(1) = 4/5$. Calculate:
   (a) $\mathrm{E}Z$
   (b) $\mathrm{E}Z^2$
   (c) $\mathrm{E}Z^3$
   (d) $\mathrm{Var}(Z)$
   (e) $\mathrm{E}e^{tZ}$, where $t \in \mathbb{R}$.

   What kind of random variable is Z (i.e. what is the name of its distribution?)

2 Discrete Probability distributions

1. Answer the following questions about a binomial random variable with parameter p:
   (a) Write down its probability mass function.
   (b) What is its mean? Its variance?
   (c) Given at least one example of a situation that can be modeled with a binomial random variable.
2. Answer the following questions about a Poisson random variable with parameter λ:
   (a) Write down its probability mass function.
   (b) What is its mean? Its variance?
   (c) Do you remember how a Poisson random variable is related to a binomial?
3. In 1898, L. J. Bartkiewicz published data showing the # of soldiers killed by horse kicks each year in each corps in the Prussian cavalry, followed a Poisson distribution with mean = 0.61.
   (a) What is the probability of at least 2 deaths in a year?
   (b) What is the probability of no deaths in 5 years?
4. Consider 25 machine parts, where a part is considered acceptable only if it passes tolerance. We sample 10 parts and find that none are defective. Using the hypergeometric distribution, what is the probability of this event if we know that there are 6 defectives among the 25 parts.
5. Suppose that in a large population of fruit flies, the proportion having vestigial wings is p. We decide to sample until we have found 100 such flies.
   (a) How long do we have to wait in expectation to get the first such fly?
   (b) What is the probability that we will have to examine exactly a total of N flies? What distribution is this?

3 Continuous distributions

1. Recall a random variable X has an *exponential* distribution with parameter λ if it has probability density function
$$f(x) = \begin{cases} Ce^{-\lambda x} & x > 0 \\ 0 & x < 0 \end{cases}$$
   for some constant C.
   (a) Determine the constant C that makes this a probability density function.
   (b) What is $P(\lambda < X < 3\lambda)$?
   (c) Find the cumulative distribution function of X.
2. Suppose X is a normal random variable with mean 2 and standard deviation 5.
   (a) Write down the pdf of X.

(b) What transformation can you do to X, to make it a standard normal?

3. The average height of a woman aged 20-74 years in the US is 64 inches, and the standard deviation is 3 inches.[1] Suppose the height of a woman is normally distributed.

   (a) What is the probability that a randomly selected woman is between 60 and 70 inches?

   (b) What is the height of a woman who is taller than exactly 95% of the population?

   (c) What is the height of a woman who is shorter than exactly 90% of the population?

   (d) What is the probability that 5 women selected at random all exceed 68 inches?

4. Review the Gamma, Beta, and Cauchy distribution.

4 Multiple random variables

1. Suppose X, Y have joint probability mass function

$X \backslash Y$	1	2	3
1	0.05	0.05	0.2
2	0.1	0.2	0.05
3	0.2	0.1	0.05

   (a) Calculate the probability mass function of $Z = X + Y$.

   (b) Calculate the covariance of X, Y.

   (c) Calculate the correlation of X, Y.

2. A *bivariate normal* is a pair of random variables (X, Y) with the following joint density function:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}-\frac{2\rho xy}{\sigma_x\sigma_y}\right)}$$

   Here $\rho \geq 0$ is the correlation between X and Y, and $\sigma_x, \sigma_y > 0$.

   (a) Compute the marginal densities $f_X(x)$, $f_Y(y)$.

   (b) Explain why each of X, Y is individually Gaussian with means 0 and variances σ_x^2, σ_y^2 respectively.

   (c) Show that X, Y are independent if and only if $\rho = 0$.

5 LLT and CLT

1. Explain the Law of Large Numbers and the Central Limit Theorem.

2. Suppose you play roulette and bet 1$ on black each time. Your net winnings on the ith play, X_i, are 1 with probability 18/38 and -1 with probability 20/38. What is the approximate probability that your net winnings are ≥ 0 after 81 plays?

[1] The actual numbers are 63.7, 2.7, but work with the rounded ones.