Stat S-110 Homework 4, Summer 2015

Due: Wednesday 7/22 at 12:01 pm (Eastern Time), submitted as a PDF file via the course website or as a hard copy in class. Show your work and explain your steps.

Collaboration Policy: You are welcome to discuss the problems with others, but *you must write up your solutions yourself and in your own words.* See the syllabus for further information about academic integrity.

The following problems are from Chapters 5-6 of the book.

1. (BH 5.1) The Rayleigh distribution from Example 5.1.7 has PDF

$$f(x) = xe^{-x^2/2}, \quad x > 0.$$

Let X have the Rayleigh distribution.

(a) Find $P(1 < X < 3)$.

(b) Find the first quartile, median, and third quartile of X; these are defined to be the values q_1, q_2, q_3 (respectively) such that $P(X \leq q_j) = j/4$ for $j = 1, 2, 3$.

2. (BH 5.4) Let X be a continuous r.v. with CDF F and PDF f.

(a) Find the conditional CDF of X given $X > a$ (where a is a constant with $P(X > a) \neq 0$). That is, find $P(X \leq x | X > a)$ for all a, in terms of F.

(b) Find the conditional PDF of X given $X > a$ (this is the derivative of the conditional CDF).

(c) Check that the conditional PDF from (b) is a valid PDF, by showing directly that it is nonnegative and integrates to 1.

3. (BH 5.5) A circle with a random radius $R \sim \text{Unif}(0, 1)$ is generated. Let A be its area.

(a) Find the mean and variance of A, without first finding the CDF or PDF of A.

(b) Find the CDF and PDF of A.

4. (BH 5.6) The 68-95-99.7% rule gives approximate probabilities of a Normal r.v. being within $1, 2$, and 3 standard deviations of its mean. Derive analogous rules for the following distributions.

(a) $\text{Unif}(0, 1)$.

(b) $\text{Expo}(1)$.

(c) Expo(1/2). Discuss whether there is one such rule that applies to all Exponential distributions, just as the 68-95-99.7% rule applies to all Normal distributions, not just to the standard Normal.

5. (BH 5.14) Let $U_1, \dots, U_n$ be i.i.d. Unif$(0, 1)$, and $X = \max(U_1, \dots, U_n)$. What is the PDF of X? What is EX?

Hint: Find the CDF of X first, by translating the event $X \leq x$ into an event involving $U_1, \dots, U_n$.

6. (BH 5.18) The *Pareto distribution* with parameter $a > 0$ has PDF $f(x) = a/x^{a+1}$ for $x \geq 1$ (and 0 otherwise). This distribution is often used in statistical modeling.

(a) Find the CDF of a Pareto r.v. with parameter a; check that it is a valid CDF.

(b) Suppose that for a simulation you want to run, you need to generate i.i.d. Pareto(a) r.v.s. You have a computer that knows how to generate i.i.d. Unif$(0, 1)$ r.v.s but does not know how to generate Pareto r.v.s. Show how to do this.

7. (BH 5.24) The distance between two points needs to be measured, in meters. The true distance between the points is 10 meters, but due to measurement error we can't measure the distance exactly. Instead, we will observe a value of $10 + \epsilon$, where the error ϵ is distributed $\mathcal{N}(0, 0.04)$. Find the probability that the observed distance is within 0.4 meters of the true distance (10 meters). Give both an exact answer in terms of Φ and an approximate numerical answer.

8. (BH 5.28) Walter and Carl both often need to travel from Location A to Location B. Walter walks, and his travel time is Normal with mean w minutes and standard deviation σ minutes (travel time can't be negative without using a tachyon beam, but assume that w is so much larger than σ that the chance of a negative travel time is negligible).

Carl drives his car, and his travel time is Normal with mean c minutes and standard deviation 2σ minutes (the standard deviation is larger for Carl due to variability in traffic conditions). Walter's travel time is independent of Carl's. On a certain day, Walter and Carl leave from Location A to Location B at the same time.

(a) Find the probability that Carl arrives first (in terms of Φ and the parameters). For this you can use the important fact, proven in the next chapter, that if X_1 and X_2 are independent with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, then $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

(b) Give a fully simplified criterion (*not* in terms of Φ), such that Carl has more than a 50% chance of arriving first if and only if the criterion is satisfied.

(c) Walter and Carl want to make it to a meeting at Location B that is scheduled to begin $w + 10$ minutes after they depart from Location A. Give a fully simplified criterion (*not* in terms of Φ) such that Carl is more likely than Walter to make it on time for the meeting if and only if the criterion is satisfied.

9. (BH 5.39) Three students are working independently on their probability homework. All 3 start at 1 pm on a certain day, and each takes an Exponential time with mean 6 hours to complete the homework. What is the earliest time when all 3 students will have completed the homework, on average? (That is, at this time all 3 students need to be done with the homework.)

10. (BH 6.9) Let Y be Log-Normal with parameters μ and σ^2. So $Y = e^X$ with $X \sim \mathcal{N}(\mu, \sigma^2)$. Evaluate and explain whether or not each of the following arguments is correct.

(a) Student A: "The median of Y is e^μ because the median of X is μ and the exponential function is continuous and strictly increasing, so the event $Y \le e^\mu$ is the same as the event $X \le \mu$."

(b) Student B: "The mode of Y is e^μ because the mode of X is μ, which corresponds to e^μ for Y since $Y = e^X$."

(c) Student C: "The mode of Y is μ because the mode of X is μ and the exponential function is continuous and strictly increasing, so maximizing the PDF of X is equivalent to maximizing the PDF of $Y = e^X$."

11. (BH 6.15) Let $W = X^2 + Y^2$, with X, Y i.i.d. $\mathcal{N}(0,1)$. The MGF of X^2 turns out to be $(1-2t)^{-1/2}$ for $t < 1/2$ (you can assume this).

(a) Find the MGF of W.

(b) What famous distribution that we have studied so far does W follow (be sure to state the parameters in addition to the name)? In fact, the distribution of W is also a special case of two more famous distributions that we will study in later chapters!

12. (BH 6.25) Let $Y = X^\beta$, with $X \sim \text{Expo}(1)$ and $\beta > 0$. The distribution of Y is called the *Weibull* distribution with parameter β. This generalizes the Exponential, allowing for non-constant hazard functions. Weibull distributions are widely used in statistics, engineering, and survival analysis; there is even an 800-page book devoted to this distribution: *The Weibull Distribution: A Handbook* by Horst Rinne.

For this problem, let $\beta = 3$.

(a) Find $P(Y > s + t | Y > s)$ for $s, t > 0$. Does Y have the memoryless property?

(b) Find the mean and variance of Y, and the nth moment $E(Y^n)$ for $n = 1, 2, \ldots$.

(c) Determine whether or not the MGF of Y exists.