STAT547C: Topics in Probability **Winter Term 1 (Fall) 2020-21**

Assignment 3

Due date: Nov. 10, 11:59 pm (PST). *Instructor: Benjamin Bloem-Reddy*

Instructions

Academic integrity policy: I encourage you to discuss verbally with other students about the assignment. However, you should write your answers by yourself. For example, copying (either manually or electronically) part of a function or a LaTeX equation is not permitted; and if you use online resources, you must cite them. If you discuss the assignment with anyone (a classmate or anyone else), you must say so at the top of your solutions. Also, refrain from looking at answer keys from other schools, previous years, or the Math Stack Exchange. For more information, see:
`http://learningcommons.ubc.ca/guide-to-academic-integrity/`

1. *Justify your answers formally.*
2. Submit your work before class on the due date to Gradescope (.pdf) and Canvas (.tex).

1 Measures from indefinite integrals

1. Recall that for a measure space $(E, \mathcal{E}, \mu)$ and a measurable function $p : E \to \mathbb{R}_+$, the indefinite integral of p with respect to μ is

$$\nu(A) = \mu(p\mathbf{1}_A) = \int_A \mu(dx)p(x) \ , \quad A \in \mathcal{E} \ .$$

Assume that $\mu p = \int_E \mu(dx)p(x) = 1$. Show that ν is a probability measure on $(E, \mathcal{E})$.

[15 mark(s)]

2. If $g : E \to \mathbb{R}_+$ is measurable, show that

$$\nu g = \int_E \nu(dx)g(x) = \int_E \mu(dx)p(x)g(x) = \mu(pg) \ .$$

Note: This is Proposition I.5.6 in Çinlar. Rather than give his proof, prove this directly, first with simple functions and then for positive measurable functions.

[10 mark(s)]

Question total: [25 mark(s)]

2 Markov's inequality

1. State Markov's inequality.

[5 mark(s)]

2. Prove the following: Let X be a $\mathbb{R}$-valued random variable, and let $h : \mathbb{R} \to [0, M]$ denote a non-negative (measurable) function taking values bounded by some number $M > 0$. Then for all $0 \le a < M$,

$$\mathbb{P}(h(X) \ge a) \ge \frac{\mathbb{E}[h(X)] - a}{M - a}.$$

[5 mark(s)]

3. Let $X_1, X_2, \dots$ be a sequence of i.i.d. $\mathbb{R}_+$-valued random variables with distribution μ, $S_n = \frac{1}{n}\sum_{i=1}^n X_i$, and let $a \in \mathbb{R}$. Let $\hat{\mu}(r) = \int_0^\infty e^{rx}\mu(dx)$ denote the Laplace transform of μ.[1] Prove that for every $t > 0$,

$$\mathbb{P}(S_n \ge a) \le e^{-ta}(\hat{\mu}(t/n))^n \ .$$

[8 mark(s)]

4. Use this to show that when the X_i's (as in the previous part above) are i.i.d. Gamma(α, β) random variables, for any $\delta > 0$,

$$\mathbb{P}(S_n \ge \mathbb{E}[X_1](1 + \delta)) \le e^{-n\alpha\delta}(1 + \delta)^{n\alpha} \ .$$

[12 mark(s)]

Question total: [30 mark(s)]

3 Simplifiable random variables and beta-gamma algebra

(a) Let X, Y, Z be random variables taking values in $\mathbb{R}_+$, with $X \perp\!\!\!\perp Y$ and $Z \perp\!\!\!\perp Y$. Y is said to be **simplifiable** if $XY \overset{d}{=} YZ$ implies $X \overset{d}{=} Z$.

Assume that the characteristic function of $\log Y$ has only isolated zeros (i.e., every zero has a neighborhood that does not contain any other zeros). (Since it is an analytic function on $\mathbb{C}$, if it had non-isolated zeros then it would be zero everywhere.)

Prove that if $\mathbb{P}(Y = 0) = 0$ then Y is simplifiable.

Warning! $\log X$ is not well-defined when $X = 0$, which the hypothesis does not rule out (likewise for $\log Z$).

[5 mark(s)]

(b) Let $\gamma_{a,b}$ be a gamma random variable with scale parameter a and rate parameter b, and $\beta_{a,b}$ a beta random variable with parameters a and b. That is,

$$\mathbb{P}(\gamma_{a,b} \in dx) = \lambda(dx)x^{a-1}e^{-bt}b^a/\Gamma(a) \ , \quad x \in \mathbb{R}_+ \ ,$$

[1] You might observe that in class we defined the Laplace transform as $\mathbb{E}[e^{-rX}]$. The difference is just a convention of sign, and using $\mathbb{E}[e^{rX}]$ is convenient for this problem.

and

$$\mathbb{P}(\beta_{a,b} \in dx) = \lambda(dx) \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} , \quad x \in [0,1] .$$

Let γ_a be a gamma random variable with scale parameter a, and $b = 1$.

We saw in class that for $\gamma_a \perp\!\!\!\perp \gamma_b$,

$$\left(\gamma_a + \gamma_b, \frac{\gamma_a}{\gamma_a + \gamma_b} \right) \stackrel{d}{=} (\gamma_{a+b}, \beta_{a,b}) , \tag{1}$$

with independence between the elements on the left-hand side, and also independence between the elements on the right-hand side.

Prove a different version of this: for $\gamma_a \perp\!\!\!\perp \gamma_b$,

$$(\gamma_a, \gamma_b) \stackrel{d}{=} (\beta_{a,b}\gamma_{a+b}, (1 - \beta_{a,b})\gamma_{a+b}) , \tag{2}$$

with $\beta_{a,b} \perp\!\!\!\perp \gamma_{a+b}$.

[5 mark(s)]

(c) Show the identity (with independence on the right-hand side)

$$\beta_{a,b+c} \stackrel{d}{=} \beta_{a,b}\beta_{a+b,c} . \tag{3}$$

Hint: Use the fact that γ_{a+b+c} is simplifiable.

[5 mark(s)]

Question total: [15 mark(s)]

4 Size bias

(a) Let X be a random variable taking values in $\mathbb{R}_+$, with $\mathbb{E}[X] \in (0, \infty)$. Let its distribution be μ. Let X^* be another random variable with distribution $\nu \ll \mu$, such that

$$\frac{d\nu}{d\mu}(x) = \frac{\mathbb{P}(X^* \in dx)}{\mathbb{P}(X \in dx)} = \frac{x}{\mathbb{E}[X]} . \tag{4}$$

X^* is said to be a *size-biased* version of X, and to have the size-biased X distribution. Assume that μ is absolutely continuous with respect to the Lebesgue measure λ on $\mathbb{R}_+$, so that X has density f_X. What is the density of X^*, f_{X^*}?

[5 mark(s)]

(b) Let X be a random variable with Poisson(α) distribution. Show that $X^* \stackrel{d}{=} X + 1$.

[5 mark(s)]

(c) Let $\gamma_{a,b}$ be a gamma random variable, as above. What is the distribution of the size-biased version $\gamma^*_{a,b}$?

[5 mark(s)]

(d) Let $\beta_{a,b}$ be a beta random variable, as above. What is the distribution of the size-biased version $\beta^*_{a,b}$?

[5 mark(s)]

(e) Let $\beta_{a,1}$ be a beta random variable, and $\beta^*_{a,1}$ a size-biased version such that $\beta_{a,1} \perp\!\!\!\perp \beta^*_{a,1}$. What is the distribution of $\beta_{a,1}\beta^*_{a,1}$?

[5 mark(s)]

Question total: [25 mark(s)]

Assignment total: [95 mark(s)]