STAT 512: Statistical Inference **Autumn 2022**

Quiz Session 4: Practice Midterm Problems

Yikun Zhang *October 26, 2022*

The following problems are partially adopted from quiz sessions developed by Zhen Miao and Apara Venkat for STAT 512 in Autumn 2021.

Problem 1 (Expectation of a Positive Random Variable)**.** *Suppose X is a positive random variable with finite expectations,* i.e., $\mathbb{E}(X) < \infty$. *Show that* $\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > x)\, dx$.

Proof. We start from the definition of $\mathbb{E}(X)$ and compute that

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}\int_0^X dt \\
&= \mathbb{E}\left(\int_0^\infty \mathbb{1}_{\{y \le X\}} dy\right) \\
&= \int_0^\infty \left(\mathbb{E}\left[\mathbb{1}_{\{y \le X\}}\right]\right) dy \\
&= \int_0^\infty \mathbb{P}(X > y) dy.
\end{aligned}$$

The result follows. □

Problem 2 (Tail Bound for a Standard Normal Distribution)**.** *Suppose Z is a standard normal random variable with $z \mapsto \phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2)$ as its density function.*

(a) Show $\phi'(z) + z \cdot \phi(z) = 0$.

(b) Use (a) to prove

$$P(Z \ge z) \le \frac{\phi(z)}{z} \text{ for all } z > 0.$$

Proof. (a) It follows from $\phi'(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2) \cdot \frac{d}{dz}(-z^2/2) = -z\phi(z)$ that $\phi'(z) + z\phi(z) = 0$.

(b) To bound $P(Z \ge z)$, note that

$$\begin{aligned}
P(Z \ge z) &= \int_z^\infty \phi(x) dx \\
&= \int_z^\infty -\frac{\phi'(x)}{x} dx \\
&= \int_z^\infty -\frac{1}{x} d\phi(x) \\
&= -\frac{\phi(x)}{x}\Bigg|_z^\infty - \int_z^\infty \frac{\phi(x)}{x^2} dx
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\phi(z)}{z} - \int_z^\infty \frac{\phi(x)}{x^2} dx \\
&\le \frac{\phi(z)}{z}.
\end{aligned}
$$

The result follows. □

Problem 3. *Let $X \sim$ Uniform$(0, \pi/2)$ and $\gamma > 0$ be a constant. Define $Y = \gamma \tan(X)$. Find the distribution of Y and $\mathbb{E}(Y)$.*

Solution. Consider the CDF of Y as

$$
\begin{aligned}
\mathbb{P}(Y \le y) &= \mathbb{P}\left(\gamma \tan(X) \le y\right) \\
&= \mathbb{P}\left(X \le \arctan\left(\frac{y}{\gamma}\right)\right) \\
&= \frac{2}{\pi} \arctan\left(\frac{y}{\gamma}\right)
\end{aligned}
$$

when $\arctan\left(\frac{y}{\gamma}\right) \in \left(0, \frac{\pi}{2}\right)$, or equivalently, $y \in (0, \infty)$. It implies that the PDF of Y is

$$
f_Y(y) = \frac{d}{dy}\mathbb{P}(Y \le y) = \frac{2}{\pi}\left(\frac{\gamma}{\gamma^2 + y^2}\right)
$$

with $y \in (0, \infty)$, which has a form as the Cauchy distribution.

Notes: you can also obtain the same answer by using Theorem 2.1 in Lecture 2 notes. □

Remark 1. *Different from the usual Cauchy distribution with $(-\infty, \infty)$, the mean/expectation of Y does exist but is infinite, because*

$$
\mathbb{E}(Y) = \int_0^\infty \frac{2}{\pi}\left(\frac{\gamma y}{\gamma^2 + y^2}\right) dy = \frac{\gamma}{\pi} \log\left(\gamma^2 + y^2\right)\Big|_0^\infty = \infty - \frac{2\gamma}{\pi}\log\gamma = \infty.
$$

However, if we assume that $X \sim$ Uniform$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $Y = \gamma \tan(X)$ with $\gamma > 0$, then the PDF of Y becomes

$$
\tilde{f}_Y(y) = \frac{1}{\pi}\left(\frac{\gamma}{\gamma^2 + y^2}\right)
$$

with $y \in (-\infty, \infty)$. In this case, Y follows the standard Cauchy distribution with scale parameter γ and its mean does not exist (or is undefined), because both integrals are infinite but with a different sign:

$$
\int_{-\infty}^0 \frac{1}{\pi}\left(\frac{\gamma y}{\gamma^2 + y^2}\right) dy = -\infty \quad \text{and} \quad \int_0^\infty \frac{1}{\pi}\left(\frac{\gamma y}{\gamma^2 + y^2}\right) dy = \infty.
$$

Nevertheless, the even-powered raw moments $\mathbb{E}|Y|^k$ with k being an even integer exist but are infinite.

Problem 4 (2017 MS Theory Exam)**.** *Let $X \sim$ Exponential(λ), $\lambda > 0$. Suppose that we only observe the fractional parts*

$$
Y = X - \lfloor X \rfloor,
$$

where $\lfloor x \rfloor$ is the largest integer that is smaller than x. For example, if $X = 5.6$, then you observe $Y = 0.6$ and if $X = 5$, then you observe $Y = 0$. Note that $Y \in [0, 1)$. What is the distribution of each Y?
Hint: What are all the possible values of X when $Y = y$?

Solution. Notice that when $Y = y$, then X should be equal to either one of $y, 1+y, 2+y, \dots$, because $y \in [0, 1)$ and $X - Y$ must be an integer.

Now, consider the CDF of Y and calculate that

$$\begin{aligned}
\mathbb{P}(Y \le y) &= P(0 \le X \le y) + P(1 \le X \le 1+y) + P(2 \le X \le 2+y) + \dots \\
&= \sum_{k=0}^{\infty} P(k \le X \le k+y) \\
&= \sum_{k=0}^{\infty} \int_k^{k+y} \lambda e^{-\lambda x} dx \\
&= \sum_{k=0}^{\infty} e^{-\lambda k} - e^{-\lambda(k+y)} \\
&= \sum_{k=0}^{\infty} e^{-\lambda k}(1 - e^{-\lambda y}) \\
&= (1 - e^{-\lambda y}) \sum_{k=0}^{\infty} e^{-\lambda k}
\end{aligned}$$

The summation is a geometric series. (Recall that $|a| < 1$, then $\sum_{r=0}^{\infty} a^r = \frac{1}{1-a}$.) Here, $a = e^{-\lambda}$. Therefore, the CDF of Y is

$$\mathbb{P}(Y \le y) = \frac{1 - e^{-\lambda y}}{1 - e^{-\lambda}},$$

whose corresponding PDF is $f_Y(y) = \frac{\lambda e^{-\lambda y}}{1 - e^{-\lambda}}$ with $y \in [0, 1)$. Thus, Y indeed follows a truncated exponential distribution to $[0, 1)$. □

Problem 5 (Poisson-Gamma)**.** *Let $X \sim$ Poisson(Λ), where $\Lambda \sim$ Gamma(α, β). Find the distribution of $\Lambda \mid X = x$ and the posterior mean $\mathbb{E}(\Lambda \mid X = x)$.*

Reminders: The PMF of $X \sim$ Poisson(λ) is $\mathbb{P}(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$, and the PDF of $\Lambda \sim$ Gamma(α, β) is $f_\Lambda(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$.

Solution. By Bayes' rule, the conditional PDF of $\Lambda | X = x$ is given by

$$\begin{aligned}
p(\lambda|x) &= \frac{p(x|\lambda) \cdot p(\lambda)}{p(x)} \\
&\propto p(x|\lambda) \cdot p(\lambda) \\
&= \frac{\lambda^x}{x!} e^{-\lambda} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \\
&\propto \lambda^{x+\alpha-1} e^{-\lambda(\beta+1)},
\end{aligned}$$

where we drop all the constants (including the factors that only depend on x) in the last step. Thus, the PDF of $\Lambda \mid x$ is of the form

$$p(\lambda \mid x) = K \cdot \lambda^{x+\alpha-1} e^{-\lambda(\beta+1)}$$

where K is the normalizing constant such that $\int_\lambda p(\lambda \mid x) d\lambda = 1$. Thus, we show that

$$\Lambda | X = x \sim \text{Gamma}(x + \alpha, 1 + \beta).$$

Finally, the posterior mean $\mathbb{E}(\Lambda \mid X = x)$ is given by

$$\begin{aligned}
\mathbb{E}(\Lambda \mid X = x) &= \int_0^\infty \lambda \cdot \frac{(1+\beta)^{\alpha+x}}{\Gamma(\alpha+x)} \lambda^{\alpha+x-1} e^{-(\beta+1)\lambda} d\lambda \\
&= \frac{\Gamma(\alpha+x+1)}{\Gamma(\alpha+x)\cdot(1+\beta)} \underbrace{\int_0^\infty \frac{(1+\beta)^{\alpha+x+1}}{\Gamma(\alpha+x+1)} \lambda^{\alpha+x+1-1} e^{-(\beta+1)\lambda} d\lambda}_{\text{PDF of Gamma}(x+\alpha+1,\beta+1)} \\
&= \frac{\alpha+x}{\beta+1}.
\end{aligned}$$

The results follow. □

Remark 2. *Notice that the prior PDF $p(\lambda)$ and the posterior PDF $p(\lambda \mid x)$ belong to the same distribution family. In this case, we call it a conjugate prior and say that, "the conjugate prior of the Poisson distribution is the gamma distribution."*

Problem 6 (Summation of Independent Gamma Random Variables)**.** *Let $X_1, \ldots, X_n$ be independently distributed variables such that $X_i \sim Gamma(\alpha_i, \beta)$. What is the distribution of $T = \sum_{i=1}^n X_i$?*

You may use the fact that the MGF of $X \sim Gamma(a, b)$ is $M_X(t) = \left(1 - \frac{t}{b}\right)^a$ where $t < b$.

Solution: An easy way to solve this is by using moment generating functions. We know the MGF of X_i is

$$\begin{aligned}
M_{X_i}(t) &= \mathbb{E}[e^{tX_i}] \\
&= \left(1 - \frac{t}{\beta}\right)^{-\alpha_i}, \quad t < \beta \qquad \text{(can you show this?)}
\end{aligned}$$

We know that all X_i are independent. Therefore,

$$\begin{aligned}
M_T(t) &= M_{\sum_{i=1}^n X_i}(t) \\
&= \mathbb{E}[e^{t\sum_{i=1}^n X_i}] \\
&= \mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right] \\
&= \prod_{i=1}^n \mathbb{E}\left[e^{tX_i}\right] \qquad \text{(independence)} \\
&= \prod_{i=1}^n \left(1 - \frac{t}{\beta}\right)^{-\alpha_i}, \quad t < \beta \\
&= \left(1 - \frac{t}{\beta}\right)^{-\sum_{i=1}^n \alpha_i}, \quad t < \beta
\end{aligned}$$

Given that MGFs uniquely determine the distribution, we conclude that $\sum_{i=1}^n X_i \sim \text{Gamma}\left(\sum_{i=1}^n \alpha_i, \beta\right)$.

■

Exercise. Can you arrive at the same result by deriving the CDFs of $\sum_{i=1}^n X_i$?
Hint: Start with $X_1 + X_2$. Then use induction to complete the argument.

Problem 7 (Acceptance-Rejection Sampling). *How can we sample $X \sim F$ with F being a distribution that has a closed-form PDF f?*

Assume that we can sample $U \sim Unif(0,1)$ infinitely many times. Also, assume that we can sample $Y \sim G$ infinitely many times, where G is a known distribution with density g whose support will include the support of F. Note that X, Y, U are all independent of each other.

Define c to be a fixed constant such that $c \geq \sup_x \frac{f(x)}{g(x)}$, where $c \in [1, \infty)$ and we generally want c to be as close as possible to 1. (To see why, look at part (a).) The algorithm is as follows

1. *Generate $Y \sim G$.*
2. *Generate $U \sim Unif(0,1)$.*
3. *If $U \leq \frac{f(Y)}{c \cdot g(Y)}$, then "accept" $X := Y$.*
4. *Else, "reject" and go back to step 1.*

(a) *What is the probability that we accept?*

(b) *Let N be the number of iterative times until an X is accepted. What is the distribution of N? On average, how many iterative times will be until we accept an X?*

(c) *Show that the accepted values indeed come from distribution F.*

Hint: What is the distribution of the accepted values in Step 3?

Solution. (a) First, we look at the conditional probability as

$$\begin{aligned}\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)} \mid Y = y\right) &= \mathbb{P}\left(U \leq \frac{f(y)}{c \cdot g(y)}\right) \\ &= \frac{f(y)}{c \cdot g(y)}\end{aligned}$$

Then, the probability of acceptance is

$$\begin{aligned}\mathbb{P}(\text{"Accept"}) &= \mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)}\right) \\ &= \int_{y=-\infty}^{\infty} \mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)} \mid Y = y\right) g(y)dy \\ &= \int_{y=-\infty}^{\infty} \frac{f(y)}{c \cdot g(y)} g(y)dy \\ &= \frac{1}{c} \int_{y=-\infty}^{\infty} f(y)dy \\ &= \frac{1}{c}\end{aligned}$$

(b) By definition, this is a geometric distribution with probability of success as $p := \frac{1}{c}$ (think about it). In particular, $\mathbb{P}(N = n) = (1-p)^{n-1}p$ where $n = 0, 1, \ldots$. The expectation of N is

$$\mathbb{E}(N) = \sum_{n=0}^{\infty} n(1-p)^{n-1}p = \frac{1}{p} = c,$$

where we recall the diagnostic exercise 2 of Quiz 1 to obtain the second equality.

(c) The distribution of accepted values is a conditional distribution, $Y \mid U \leq \frac{f(Y)}{c \cdot g(Y)}$. Let us show that this is the same as F.

$$\begin{aligned}
\mathbb{P}\left(Y \leq y \mid U \leq \frac{f(Y)}{c \cdot g(Y)}\right) &= \frac{\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)} \mid Y \leq y\right) \cdot \mathbb{P}\left(Y \leq y\right)}{\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)}\right)} && \text{(Bayes' rule)} \\
&= \frac{\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)} \mid Y \leq y\right) \cdot G(y)}{\frac{1}{c}}
\end{aligned}$$

The probability in the numerator is slightly tricky.

$$\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)} \mid Y \leq y\right) = \frac{\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)}, Y \leq y\right)}{P(Y \leq y)}$$

and

$$\begin{aligned}
\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)}, Y \leq y\right) &= \int_{-\infty}^{\infty} \mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)}, Y \leq y \middle| Y = t\right) g(t)dt \\
&= \int_{-\infty}^{\infty} \mathbb{P}\left(U \leq \frac{f(y)}{c \cdot g(y)}, t \leq y\right) g(t)dt \\
&= \int_{-\infty}^{\infty} \mathbb{P}\left(U \leq \frac{f(y)}{c \cdot g(y)}\right) \cdot \mathbb{1}_{\{t \leq y\}} \cdot g(t)dt \\
&= \int_{-\infty}^{y} \mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)}\right) g(t)dt \\
&= \int_{-\infty}^{y} \frac{f(t)}{c \cdot g(t)} g(t)dt \\
&= \frac{1}{c} \int_{-\infty}^{y} f(t)dt \\
&= \frac{F(y)}{c},
\end{aligned}$$

which implies that $\mathbb{P}\left(U \leq \frac{f(Y)}{c \cdot g(Y)} \mid Y \leq y\right) = \frac{F(y)}{c \cdot G(y)}$. Finally,

$$\begin{aligned}
\mathbb{P}\left(Y \leq y \mid U \leq \frac{f(Y)}{c \cdot g(Y)}\right) &= \frac{\frac{F(y)}{c \cdot G(y)} \cdot G(y)}{\frac{1}{c}} \\
&= F(y)
\end{aligned}$$

which was the target distribution. □

Remark 3. *The interested readers can be referred to* `http://www.columbia.edu/~ks20/4703-Sigman/4703-07-Notes-ARM.pdf` *for more discussion about acceptance-rejection methods.*