# STAT 30400 : Distribution Theory

## *Fall 2019*

•

## Homework 4

•

*Solutions by*

Jinhong Du

12243476

1. (13 pts) Let $X$ be a random variable.

   (a) Show that $X$ is integrable if and only if
$$\sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n) < \infty.$$

*Proof.* Since
$$\begin{aligned} X \text{ is integrable} \quad &\iff \quad \mathbb{E}X^+ < \infty, \mathbb{E}X^- < \infty \\ &\iff \quad \mathbb{E}|X| = \mathbb{E}X^+ + \mathbb{E}X^- < \infty, \end{aligned}$$
next we just need to prove that $|X|$ is integrable if and only if $\sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n) < \infty$.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space and $E_k = \{\omega : k \leq |X(\omega)| < k+1\}$, then $\Omega = \bigcup_{k=0}^{\infty} E_k$.
Since for all $k \in \mathbb{Z}$,
$$k\mathbb{P}(E_k) \leq \int_{E_k} |X(\omega)| \mathrm{d}\mathbb{P}(\omega) \leq (k+1)\mathbb{P}(E_k),$$
we have
$$\begin{aligned} \sum_{k=0}^{\infty} k\mathbb{P}(E_k) \leq \int_{\Omega} |X(\omega)| \mathrm{d}\mathbb{P}(\omega) &\leq \sum_{k=0}^{\infty} (k+1)\mathbb{P}(E_k) \\ &\leq \sum_{k=0}^{\infty} k\mathbb{P}(E_k) + \sum_{k=0}^{\infty} \mathbb{P}(E_k) \\ &= \sum_{k=0}^{\infty} k\mathbb{P}(E_k) + 1. \end{aligned}$$
Therefore,
$$|X| \text{ is integrable if and only if } \sum_{k=0}^{\infty} k\mathbb{P}(E_k) < \infty. \tag{1}$$
$\Longrightarrow$
Since $|X|$ is integrable, $\mathbb{P}(|X| = \infty) = 0$ (Otherwise, $\mathbb{E}|X| \geq \infty \times \mathbb{P}(|X| = \infty) = \infty$). Then
$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{P}(|X| \geq n) &= \sum_{n=0}^{\infty} \sum_{k=n}^{\infty} \mathbb{P}(E_k) \\ &= \sum_{k=0}^{\infty} \sum_{n=0}^{k-1} \mathbb{P}(E_k) \\ &= \sum_{k=0}^{\infty} k\mathbb{P}(E_k). \end{aligned}$$
From (1), we have $\sum_{k=1}^{\infty} k\mathbb{P}(E_k) \leq \sum_{k=0}^{\infty} k\mathbb{P}(E_k) < \infty$.
$\Longleftarrow$
Since $\sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n) < \infty$ implies $\mathbb{P}(|X| = \infty) = 0$, otherwise $\sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n) \geq \sum_{n=1}^{\infty} \mathbb{P}(|X| = \infty) = \infty$, which is a contradiction. So again
$$\sum_{k=0}^{\infty} k\mathbb{P}(E_k) = \sum_{n=0}^{\infty} \mathbb{P}(|X| \geq n) = \mathbb{P}(|X| \geq 0) + \sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n) < \infty$$
and from (1), $|X|$ is integrable. $\square$

(b) Show that there exists a transformation $f : [0, \infty) \to [0, \infty)$ that is increasing, such that $f(0) = 0$, $f(x) \to \infty$ as $x \to \infty$, and $f(|X|)$ is integrable.

*Proof.* If $X$ is integrable, then $f(x) = x$ satisfies the conditions.
If $X$ is not integrable, $\mathbb{E}|X| = \mathbb{E}X^+ + \mathbb{E}X^- = \infty$.
Let $Q(u) = \inf\{x : F_{|X|}(x) \geq u\}$ be the quantile function of $|X|$. Define

$$f(x) = \begin{cases} 0, & 0 \\ n, & Q\left(1 - \frac{1}{2^{n-1}}\right) < x \leq Q\left(1 - \frac{1}{2^n}\right) \qquad n = 1, 2, \ldots \end{cases}$$

Obviously, $f(0) = 0$ and $f$ is increasing since $Q(u)$ is non-decreasing.
If $F_{|X|}(x_0) = 1$ for some $x_0 \in \mathbb{R}$, then $\mathbb{E}|X| \leq x_0 < \infty$, i.e., $X$ is integrable. So if $X$ is not integrable, then $F_{|X|}(x) < 1$ for $x < \infty$. So as $n \to \infty$, $Q(1 - \frac{1}{2^n}) \to Q(1-) = \infty$. Since $\bigcup_{n=1}^{\infty}(Q\left(1 - \frac{1}{2^{n-1}}\right), Q\left(1 - \frac{1}{2^n}\right)] = (0, \infty)$, $\lim\limits_{x\to\infty} f(x) = \lim\limits_{n\to\infty} n = \infty$.
Also,

$$\begin{aligned}
\mathbb{E}f(|X|) &= \sum_{n=1}^{\infty} n\mathbb{P}\left(Q\left(1 - \frac{1}{2^{n-1}}\right) < |X| \leq Q\left(1 - \frac{1}{2^n}\right)\right) \\
&= \sum_{n=1}^{\infty} n\mathbb{P}(1 - \frac{1}{2^{n-1}} < F(|X|) \leq 1 - \frac{1}{2^n}) \\
&= \sum_{n=1}^{\infty} \frac{n}{2^{n-1}} \\
&= \sum_{n=1}^{\infty} \left.\frac{\mathrm{d}x^n}{\mathrm{d}x}\right|_{x=\frac{1}{2}} \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=1}^{\infty} x^n\right)\right|_{x=\frac{1}{2}} \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}x}\frac{x}{1-x}\right|_{x=\frac{1}{2}} \\
&= \left.\frac{1}{(1-x)^2}\right|_{x=\frac{1}{2}} \\
&= 4,
\end{aligned}$$

i.e., $f(|X|)$ is integrable. $\square$

2. (12 pts) Let $X$ and $Y$ be random variables with finite variances. We denote with $\mu_X$ and $\mu_Y$ the means, $\sigma_X$ and $\sigma_Y$ the standard deviations and with $\rho$ the correlation.

(a) Find the value of $\beta \in \mathbb{R}$ that minimizes $Var(Y - \beta X)$.

If $\sigma_X = 0$, then $X = c$ a.s.. So $Var(Y - \beta X) = Var(Y)$ is irrelevant to $\beta$.
If $\sigma_X \neq 0$,

$$\begin{aligned} Var(Y - \beta X) &= Var(Y) + \beta^2 Var(X) - 2Cov(Y, \beta X) \\ &= Var(Y) + \beta^2 Var(X) - 2\beta Cov(Y, X) \\ &= \sigma_X^2 \beta^2 - 2\rho\sigma_X\sigma_Y\beta + \sigma_Y^2 \\ &= \sigma_X^2 \left(\beta - \frac{\rho\sigma_Y}{\sigma_X}\right)^2 + (1-\rho^2)\sigma_Y^2 \\ &\geq (1-\rho^2)\sigma_Y^2, \end{aligned}$$

the inequality holds when $\beta = \frac{\rho\sigma_Y}{\sigma_X}$.

(b) Find the values of $\beta \in \mathbb{R}$ such that $Y$ and $Y - \beta X$ are uncorrelated.

If $\sigma_X = \sigma_Y = 0$ then $\rho$ does not exist. So either one must be nonzero.
If $Y$ and $Y - \beta X$ are uncorrelated, then $Cor(Y, Y - \beta X) = 0$ and $Cov(Y, Y - \beta X) = 0$. So

$$\begin{aligned} Cov(Y, Y - \beta X) &= Var(Y) - \beta Cov(Y, X) \\ &= \sigma_Y^2 - \beta\rho\sigma_X\sigma_Y \\ &= 0, \end{aligned}$$

which implies $\beta = \frac{\sigma_Y}{\rho\sigma_X}$ when $\sigma_X, \sigma_Y \neq 0$. If $\sigma_X \neq 0$ and $\sigma_Y = 0$, then $\forall\ \beta \in \mathbb{R}$ satisfies the condition. If $\sigma_X = 0$ and $\sigma_Y \neq 0$, then there is no such $\beta$.

(c) Find the values of $\beta \in \mathbb{R}$ such that $X$ and $Y - \beta X$ are uncorrelated.

Analogously,

$$\begin{aligned} Cov(X, Y - \beta X) &= Cov(X, Y) - \beta Var(X) \\ &= \rho\sigma_X\sigma_Y - \beta\sigma_X^2 \\ &= 0, \end{aligned}$$

which implies $\beta = \frac{\rho\sigma_Y}{\sigma_X}$ when $\sigma_X \neq 0$. If $\sigma_X = 0$, then $\forall\ \beta \in \mathbb{R}$ satisfies the condition.

(d) Find conditions under which, for some $\beta$, $Y - \beta X$ is uncorrelated with both $X$ and $Y$.

If $\sigma_X = 0$ and $\sigma_Y \neq 0$, then $Cov(Y - \beta X) = Var(Y) > 0$. So there is no such $\beta$.
If $\sigma_Y = 0$ and $\sigma_X \neq 0$, then $Cov(Y - \beta X) = \beta^2 Var(X)$. If $b = 0$, then $Y - \beta X$ is uncorrelated with both $X$ and $Y$.
If $\sigma_Y \neq 0$ and $\sigma_X \neq 0$, from (b) and (c), we have $Y - \beta X$ is uncorrelated with both $X$ and $Y$ when $\beta = \frac{\sigma_Y}{\rho\sigma_X} = \frac{\rho\sigma_Y}{\sigma_X}$. So $(1-\rho^2)\sigma_X\sigma_Y = 0$. Since $\sigma_Y \neq 0$ and $\sigma_X \neq 0$, we have $\rho = \pm 1$.

3. (10 pts) Let $X$ be a random variable with mean $\mu$ and variance $\sigma^2$.

(a) Show that,

$$\mathbb{P}(X - \mu \geq \alpha) \leq \frac{\sigma^2}{\sigma^2 + \alpha^2}, \qquad \alpha \geq 0.$$

*Proof.* If $\alpha = 0$, then $\mathbb{P}(X - \mu \geq \alpha) \leq 1$ always holds.
If $\alpha > 0$, $\forall\ u \geq 0$,

$$\begin{aligned}
\mathbb{P}(X - \mu \geq \alpha) &= \mathbb{P}(X - \mu + u \geq \alpha + u) \\
&\leq \mathbb{P}((X - \mu + u)^2 \geq (\alpha + u)^2) \\
&\leq \frac{\mathbb{E}[(X - \mu + u)^2]}{(\alpha + u)^2} \\
&= \frac{\sigma^2 + u^2}{(\alpha + u)^2}
\end{aligned}$$

Let $f(u) = \frac{\sigma^2 + u^2}{(\alpha + u)^2}$. By seting

$$\begin{aligned}
f'(u) &= \frac{2u(\alpha + u)^2 - 2(\alpha + u)(\sigma^2 + u^2)}{(\alpha + u)^4} \\
&= \frac{2u(\alpha + u) - 2(\sigma^2 + u^2)}{(\alpha + u)^3} \\
&= \frac{2u\alpha - 2\sigma^2}{(\alpha + u)^3} \\
&= 0,
\end{aligned}$$

we have $u^* = \frac{\sigma^2}{\alpha}$ is a stationary point. $f'(u) < 0$ when $0 \leq u < u^*$; $f'(u) > 0$ when $u > u^*$. So $f(u) \geq f(u^*) = \frac{\sigma^2}{\sigma^2 + \alpha^2}$. Let $u = u^*$, we have

$$\mathbb{P}(X - \mu \geq \alpha) \leq \frac{\sigma^2}{\sigma^2 + \alpha^2}, \qquad \alpha \geq 0.$$

□

(b) Show that,

$$\mathbb{P}(|X - \mu| \geq \alpha) \leq \frac{2\sigma^2}{\sigma^2 + \alpha^2}, \qquad \alpha \geq 0.$$

When is this better than Chebyshev's inequality?

If $\alpha = 0$, the inequality holds naturally.
If $\alpha < 0$, $\forall\ u \geq 0$,

$$\begin{aligned}
\mathbb{P}(X - \mu \leq \alpha) &= \mathbb{P}(\mu - X + u \geq -\alpha + u) \\
&\leq \mathbb{P}((\mu - X + u)^2 \leq (-\alpha + u)^2) \\
&\leq \frac{\mathbb{E}[(\mu - X + u)^2]}{(-\alpha + u)^2} \\
&= \frac{\sigma^2 + u^2}{(-\alpha + u)^2}
\end{aligned}$$

Solution (cont.)

Analogously, let $u = u^* = \frac{\sigma^2}{-\alpha}$, we have

$$\mathbb{P}(X - \mu < \alpha) \leq \frac{\sigma^2}{\sigma^2 + \alpha^2}, \qquad \alpha < 0.$$

Therefore, for $\alpha \geq 0$,

$$\begin{aligned}
\mathbb{P}(|X - \mu| \geq \alpha) &= \mathbb{P}(X - \mu \geq \alpha) + \mathbb{P}(\mu - X \geq \alpha) \\
&= \mathbb{P}(X - \mu \geq \alpha) + \mathbb{P}(X - \mu \leq -\alpha) \\
&\leq \frac{2\sigma^2}{\sigma^2 + \alpha^2}.
\end{aligned}$$

For the Chebyshev's inequality,

$$\mathbb{P}(|X - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha}$$

$\frac{2\sigma^2}{\sigma^2+\alpha^2} \leq \frac{\sigma^2}{\alpha}$ when $\sigma^2 \geq 1$, or, $\sigma^2 < 1$ and $\alpha \in (0, 1 - \sqrt{1-\sigma^2}) \cup (1 + \sqrt{1-\sigma^2}, \infty)$.

4. (15 pts) Let $X_r$ denote a Gamma$(r, 1)$ random variable.

(a) Find the $g$-means of $X_r$ for the power transformations defined as $g_\lambda(x) = x^\lambda$, $\lambda \neq 0$ and $g_0(x) = \log(x)$.

Since $X_r \sim \text{Gamma}(r, 1)$, the density function of $X_r$ is given by

$$f_{X_r}(x) = \frac{1}{\Gamma(r)} x^{r-1} e^{-x} \mathbb{1}_{(0,\infty)}.$$

Then

$$\begin{aligned}
\mathbb{E} g_\lambda(X_r) &= \int_0^\infty x^\lambda \frac{1}{\Gamma(r)} x^{r-1} e^{-x} \mathrm{d}x \\
&= \frac{\Gamma(r+\lambda)}{\Gamma(r)} \int_0^\infty \frac{1}{\Gamma(r+\lambda)} x^{\lambda+r-1} e^{-x} \mathrm{d}x \\
&= \frac{\Gamma(r+\lambda)}{\Gamma(r)}.
\end{aligned}$$

So

$$g_\lambda^{-1}(\mathbb{E} g_\lambda(X_r)) = \left( \frac{\Gamma(r+\lambda)}{\Gamma(r)} \right)^{-\lambda}$$

The distribution of $g_0(X_r)$ is given by

$$\begin{aligned}
f_{g_0(X_r)}(y) &= f_{X_r}(e^y) e^{-y} \\
&= \frac{1}{\Gamma(r)} e^{(r-1)y} e^{-e^{-y}} \mathbb{1}_{(1,\infty)}
\end{aligned}$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}r} e^{(r-1)y} e^{-e^{-y}} = y e^{(r-1)y} e^{-e^{-y}} = \Gamma(r) y f_{g_0(X_r)}(y),$$

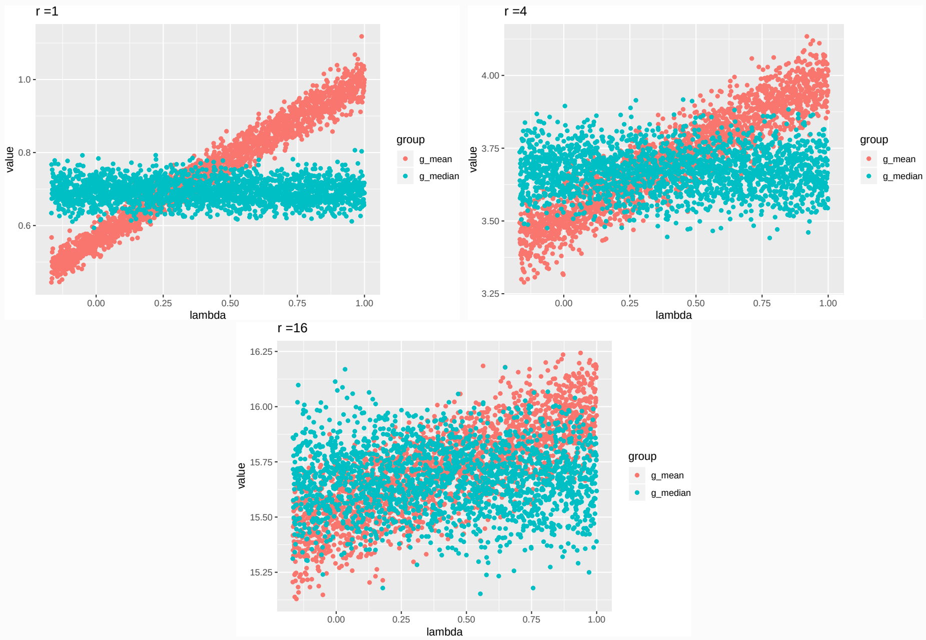
we have

$$\begin{aligned}
\mathbb{E} g_0(X_r) &= \int_{\mathbb{R}} y f_{g_0(X_r)}(y) \mathrm{d}y \\
&= \int_1^\infty \frac{1}{\Gamma(r)} \frac{\mathrm{d}}{\mathrm{d}r} e^{(r-1)y} e^{-e^{-y}} \mathrm{d}y \\
&= \frac{1}{\Gamma(r)} \frac{\mathrm{d}}{\mathrm{d}r} \int_1^\infty e^{(r-1)y} e^{-e^{-y}} \mathrm{d}y \\
&= \frac{1}{\Gamma(r)} \frac{\mathrm{d}}{\mathrm{d}r} \Gamma(r) \\
&= \frac{\mathrm{d}}{\mathrm{d}r} \log(\Gamma(r)),
\end{aligned}$$

so

$$g_0^{-1}(\mathbb{E} g_0(X_r)) = e^{\frac{\mathrm{d}}{\mathrm{d}r} \log(\Gamma(r))}.$$

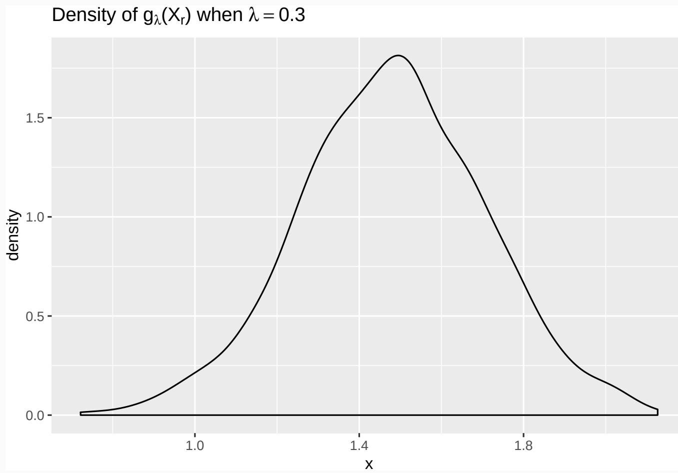
(b) For $r = 1, 4$, and 16, numerically plot and evaluate the $g_\lambda$-mean and the $g_\lambda$-median (defined similarly to the $g$-mean, as discussed in class) against $\lambda$ in the interval [-1/6, 1]. Make a separate plot for each $r$ but include the mean and median on the same plot.




```
library(ggplot2)
set.seed(1)
n <- 1000
lambda_list <- seq(-1/6,1,length.out=1000)
make_plot <- function(r){
    g_mean <- c()
    g_median <- c()
    for (i in c(1:1000)) {
        X <- rgamma(n, r)
        lambda <- lambda_list[i]
        g_mean[i] <- mean(X^lambda)^(1/lambda)
        g_median[i] <- median(X^lambda)^(1/lambda)
    }
    df1 <- data.frame(X=lambda_list, Y=g_mean, group='g_mean')
    df2 <- data.frame(X=lambda_list, Y=g_median, group='g_median')
    df <- rbind(df1,df2)
    ggplot(df, aes(x=X, y=Y, group=group,
                   color=group)) + geom_point() +
        ggtitle(sprintf('r =%d',r)) +
        xlab('lambda') + ylab('value')
}
make_plot(1)
make_plot(4)
make_plot(16)
```

(c) For which value of $\lambda$ do you think the distribution of $g_\lambda(X)$ is most nearly symmetric?

Since the density of $X_r$ is unimodal and $g_\lambda$ is strictly monotone, the density of $g_\lambda(X_r)$ is also unimode. Then if $g_\lambda(X_r)$ has equal mean and median, which equals to $g_\lambda$-mean= $g_\lambda$-median, then the density of it is mostly symmetric. In the plots, we see that $\lambda \approx 0.3$, $g_\lambda$-mean= $g_\lambda$-median, which means that the density of $g_\lambda(X_r)$ is mostly symmetric.




```r
library(latex2exp)
X <- rgamma(n, 4)
lambda <- 0.3
g_X <- X^lambda
df <- data.frame(x=g_X)
ggplot(df,aes(x=x)) +  geom_density() +
    ggtitle(TeX('Density of $g_{\\lambda}(X_{r})$ when $\\lambda = 0.3$'))
```