Homework 9

Mengxiang Jiang
Stat 610 Distribution Theory

November 10, 2025

Problem 1. The joint mgf of a random pair (X, Y) is

$$M(s,t) = \mathsf{E}(e^{sX+tY}),$$

for all (s, t) such that the expectation is finite, which we assume includes all (s, t) near $(0, 0)$.

(a) Explain why the marginal mgfs are $M_X(s) = M(s, 0)$ and $M_Y(t) = M(0, t)$.

(b) Show that

$$\mathsf{E}(XY) = \left.\frac{\partial^2 M(s,t)}{\partial s \partial t}\right|_{s=t=0}$$

and

$$\mathsf{Cov}(X,Y) = \left.\frac{\partial^2 \log(M(s,t))}{\partial s \partial t}\right|_{s=t=0}.$$

(You may assume interchangeability of derivatives and expectation.)

(c) Find the joint mgf for Example 4.5 in the notes and use it to compute $\mathsf{E}(X)$, $\mathsf{E}(Y)$, and $\mathsf{E}(XY)$. Compare to the solutions in the notes.

(a) The marginal mgf of X is defined as

$$M_X(s) = \mathsf{E}(e^{sX}) = \mathsf{E}(e^{sX+0\cdot Y}) = M(s, 0).$$

Similarly, the marginal mgf of Y is

$$M_Y(t) = \mathsf{E}(e^{tY}) = \mathsf{E}(e^{0\cdot X+tY}) = M(0, t).$$

(b) We have

$$\begin{aligned}\frac{\partial^2 M(s,t)}{\partial s \partial t} &= \frac{\partial}{\partial s}\mathsf{E}(Ye^{sX+tY})\\ &= \mathsf{E}(XYe^{sX+tY}).\end{aligned}$$

Evaluating at $s = t = 0$ gives

$$\mathsf{E}(XYe^0) = E(XY).$$

Next, we have

$$\begin{aligned}
\frac{\partial^2 \log(M(s,t))}{\partial s \partial t} &= \frac{\partial}{\partial s}(\mathsf{Var}(Y)e^{sX+tY}/M(s,t)) \\
&= \frac{\mathsf{E}(XYe^{sX+tY})M(s,t) - \mathsf{E}(Xe^{sX+tY})\mathsf{E}(Ye^{sX+tY})}{(M(s,t))^2} \\
&= \frac{\mathsf{E}(XYe^{sX+tY})}{M(s,t)} - \frac{\mathsf{E}(Xe^{sX+tY})}{M(s,t)} \cdot \frac{\mathsf{E}(Ye^{sX+tY})}{M(s,t)}.
\end{aligned}$$

Evaluating at $s = t = 0$ gives

$$\frac{E(XYe^0)}{1} - \frac{E(Xe^0)}{1} \cdot \frac{E(Ye^0)}{1} = \mathsf{Cov}(X,Y).$$

(c) Example 4.5 Suppose (X,Y) has joint pdf

$$f_{X,Y}(x,y) = \frac{1}{2}\left(\lambda^2 e^{-\lambda(x+y)} + \mu^2 e^{-\mu(x+y)}\right) 1_{(0,\infty)}(x) 1_{(0,\infty)}(y).$$

The joint mgf is

$$\begin{aligned}
M(s,t) &= \int_0^\infty \int_0^\infty e^{sx+ty}\frac{1}{2}\left(\lambda^2 e^{-\lambda(x+y)} + \mu^2 e^{-\mu(x+y)}\right) dxdy \\
&= \frac{1}{2}\left(\int_0^\infty \lambda^2 e^{-(\lambda-s)x}dx \int_0^\infty e^{-(\lambda-t)y}dy + \int_0^\infty \mu^2 e^{-(\mu-s)x}dx \int_0^\infty e^{-(\mu-t)y}dy\right) \\
&= \frac{1}{2}\left(\frac{\lambda^2}{(\lambda-s)(\lambda-t)} + \frac{\mu^2}{(\mu-s)(\mu-t)}\right), \quad s < \min(\lambda,\mu), t < \min(\lambda,\mu).
\end{aligned}$$

Using the joint mgf, we have

$$\mathsf{E}(X) = \mathsf{E}(Y) = \left.\frac{\partial M(s,t)}{\partial s}\right|_{s=t=0} = \frac{1}{2}\left(\frac{\lambda^2}{\lambda^2}\cdot\frac{1}{\lambda} + \frac{\mu^2}{\mu^2}\cdot\frac{1}{\mu}\right) = \frac{1}{2}\left(\frac{1}{\lambda} + \frac{1}{\mu}\right).$$

$$\mathsf{E}(XY) = \left.\frac{\partial^2 M(s,t)}{\partial s \partial t}\right|_{s=t=0} = \frac{1}{2}\left(\frac{\lambda^2}{\lambda^2}\cdot\frac{1}{\lambda^2} + \frac{\mu^2}{\mu^2}\cdot\frac{1}{\mu^2}\right) = \frac{1}{2}\left(\frac{1}{\lambda^2} + \frac{1}{\mu^2}\right).$$

Problem 2. Consider random pair (V, W) with joint pdf

$$f_{V,W}(v,w) = 4v^3 e^{-v(w+2)}, \quad v \geq 0, w \geq 0.$$

(a) Determine the conditional distribution for W, given V, and use it to find the best overall predictor $\mathsf{E}(W|V)$. (Take note of the marginal distribution of V for use below.)

(b) Also compute the mean squared prediction error $\mathsf{E}((\mathsf{E}(W|V) - W)^2)$.

(c) Now find $\mathsf{E}(V)$, $\mathsf{E}(W)$, $\mathsf{Var}(V)$ and $\mathsf{Cov}(V, W)$ and use them to get the best linear predictor of W from V. (Recall Theorem 4.35 in the notes.) It may help to take advantage of iterated expectation to find $\mathsf{E}(W)$ and $\mathsf{E}(VW)$.

(d) Determine the mean squared prediction error for the linear predictor and compare that to the result for part (b).

(e) In this case, the linear predictor has a serious drawback for large V. Explain. (You might want to graph the two functions for a range of values of V.)

(a) The marginal pdf of V is

$$\begin{aligned}
f_V(v) &= \int_0^\infty 4v^3 e^{-v(w+2)} dw \\
&= 4v^3 e^{-2v} \int_0^\infty e^{-vw} dw \\
&= 4v^3 e^{-2v} \cdot \frac{1}{v} = 4v^2 e^{-2v}, \quad v \geq 0.
\end{aligned}$$

The conditional pdf of W given V is

$$\begin{aligned}
f_{W|V}(w|v) &= \frac{f_{V,W}(v,w)}{f_V(v)} \\
&= \frac{4v^3 e^{-v(w+2)}}{4v^2 e^{-2v}} = ve^{-vw}, \quad w \geq 0.
\end{aligned}$$

Since this is the pdf of an exponential distribution with parameter $\frac{1}{v}$, we have

$$\mathsf{E}(W|V) = \frac{1}{v}.$$

(b) We have

$$\begin{aligned}
\mathsf{E}((\mathsf{E}(W|V) - W)^2) &= \mathsf{E}\left(\left(\frac{1}{V} - W\right)^2\right) \\
&= \mathsf{E}\left(\mathsf{E}\left(\left(\frac{1}{V} - W\right)^2 |V\right)\right) \\
&= \mathsf{E}\left(\mathsf{Var}(W|V) + (\mathsf{E}(W|V) - \mathsf{E}(W|V))^2\right) \\
&= \mathsf{E}(\mathsf{Var}(W|V)).
\end{aligned}$$

Since this is an exponential distribution with parameter $\frac{1}{v}$, we have

$$\mathsf{Var}(W|V) = \frac{1}{v^2}.$$

Therefore,

$$\mathsf{E}((\mathsf{E}(W|V) - W)^2) = \mathsf{E}\left(\frac{1}{V^2}\right).$$

To compute $\mathsf{E}\left(\frac{1}{V^2}\right)$, we have

$$\begin{aligned}\mathsf{E}\left(\frac{1}{V^2}\right) &= \int_0^\infty \frac{1}{v^2} \cdot 4v^2 e^{-2v} dv \\ &= \int_0^\infty 4e^{-2v} dv = 4 \cdot \frac{1}{2} = 2.\end{aligned}$$

(c) Since $f_V(v)$ is the pdf of a Gamma distribution with parameters $\alpha = 3$ and $\beta = \frac{1}{2}$, we have

$$\mathsf{E}(V) = \alpha\beta = 3 \cdot \frac{1}{2} = \frac{3}{2},$$

$$\mathsf{Var}(V) = \alpha\beta^2 = 3 \cdot \left(\frac{1}{2}\right)^2 = \frac{3}{4}.$$

Next, we compute $\mathsf{E}(W)$ using iterated expectation:

$$\begin{aligned}\mathsf{E}(W) = \mathsf{E}(\mathsf{E}(W|V)) &= \mathsf{E}\left(\frac{1}{V}\right) \\ &= \int_0^\infty \frac{1}{v} \cdot 4v^2 e^{-2v} dv \\ &= \int_0^\infty 4ve^{-2v} dv = 4 \cdot \frac{1}{4} = 1.\end{aligned}$$

Finally, we compute $\mathsf{E}(VW)$ using iterated expectation:

$$\mathsf{E}(VW) = \mathsf{E}(\mathsf{E}(VW|V)) = \mathsf{E}\left(V\mathsf{E}(W|V)\right) = \mathsf{E}(V \cdot \frac{1}{V}) = \mathsf{E}(1) = 1.$$

Therefore,

$$\mathsf{Cov}(V, W) = \mathsf{E}(VW) - \mathsf{E}(V)\mathsf{E}(W) = 1 - \frac{3}{2} \cdot 1 = -\frac{1}{2}.$$

The best linear predictor of W from V is

$$\mathsf{E}^*(W|V) = \mathsf{E}(W) + \frac{\mathsf{Cov}(V, W)}{\mathsf{Var}(V)}(V - \mathsf{E}(V)) = 1 + \left(-\frac{1/2}{3/4}\right)(V - \frac{3}{2}) = 2 - \frac{2}{3}V.$$

(d) The mean squared prediction error for the linear predictor is

$$\begin{aligned}\mathsf{E}((\mathsf{E}^*(W|V) - W)^2) &= \mathsf{Var}(W) - \frac{(\mathsf{Cov}(V, W))^2}{\mathsf{Var}(V)} \\ &= \mathsf{Var}(W) - \frac{(-1/2)^2}{3/4} = \mathsf{Var}(W) - \frac{1}{3}.\end{aligned}$$

To compute $\mathsf{Var}(W)$, we first compute $\mathsf{E}(W^2)$ using iterated expectation:

$$\begin{aligned}\mathsf{E}(W^2) &= \mathsf{E}(\mathsf{E}(W^2|V)) = \mathsf{E}(\mathsf{Var}(W|V) + (\mathsf{E}(W|V))^2) \\ &= \mathsf{E}\left(\frac{1}{V^2} + \left(\frac{1}{V}\right)^2\right) = 2\mathsf{E}\left(\frac{1}{V^2}\right) = 2 \cdot 2 = 4.\end{aligned}$$

Therefore,

$$\mathsf{Var}(W) = \mathsf{E}(W^2) - (\mathsf{E}(W))^2 = 4 - 1^2 = 3.$$

Thus, the mean squared prediction error for the linear predictor is

$$\mathsf{E}((\mathsf{E}^*(W|V) - W)^2) = 3 - \frac{1}{3} = \frac{8}{3} \approx 2.67.$$

This is more than the result for part (b), which was 2.

(e) The linear predictor $\mathsf{E}^*(W|V) = 2 - \frac{2}{3}V$ becomes negative for large values of V, which is not appropriate since W is a non-negative random variable.

Problem 3. Assume $U \sim \text{gamma}(2, \beta)$ and $T \sim \text{gamma}(3, \beta)$, independent.

(a) Determine the mean and variance of $V = 3U + 2T$.

(b) Determine $\mathsf{Cov}(V, W)$ where $W = 2U - 3T$. Use Theorem 4.34.i in the notes (twice).

(a) Since $U \sim \text{gamma}(2, \beta)$, we have

$$\mathsf{E}(U) = 2\beta, \quad \mathsf{Var}(U) = 2\beta^2.$$

Since $T \sim \text{gamma}(3, \beta)$, we have

$$\mathsf{E}(T) = 3\beta, \quad \mathsf{Var}(T) = 3\beta^2.$$

Therefore,

$$\mathsf{E}(V) = \mathsf{E}(3U + 2T) = 3\mathsf{E}(U) + 2\mathsf{E}(T) = 3(2\beta) + 2(3\beta) = 12\beta,$$

$$\mathsf{Var}(V) = \mathsf{Var}(3U + 2T) = 9\mathsf{Var}(U) + 4\mathsf{Var}(T) = 9(2\beta^2) + 4(3\beta^2) = 30\beta^2.$$

(b) **Theorem 4.34.** Assume X, Y and W are jointly distributed and their variances exist. i. $\mathsf{Cov}(X + W, Y) = \mathsf{Cov}(X, Y) + \mathsf{Cov}(W, Y)$.

Thus,

$$\begin{aligned}\mathsf{Cov}(V, W) &= \mathsf{Cov}(3U + 2T, 2U - 3T) \\ &= 3\mathsf{Cov}(U, 2U - 3T) + 2\mathsf{Cov}(T, 2U - 3T) \\ &= 6\mathsf{Cov}(U, U) - 9\mathsf{Cov}(U, T) + 4\mathsf{Cov}(T, U) - 6\mathsf{Cov}(T, T).\end{aligned}$$

Since U and T are independent, we have $\mathsf{Cov}(U, T) = \mathsf{Cov}(T, U) = 0$. Therefore,

$$\mathsf{Cov}(V, W) = 6\mathsf{Var}(U) - 6\mathsf{Var}(T) = 6(2\beta^2) - 6(3\beta^2) = -6\beta^2.$$

Problem 4. Suppose $(X_1, X_2) \sim$ bivariate normal$(0, 0, 1, 1, \rho)$. Let $Y_1 = X_1 + X_2 + 1$, $Y_2 = 2X_1 - X_2 - 3$. We know that the joint distribution of (Y_1, Y_2) is also bivariate normal. Find each of the five parameter values for that distribution.

Since $(X_1, X_2) \sim$ bivariate normal$(0, 0, 1, 1, \rho)$, we have

$$\mathsf{E}(X_1) = 0, \quad \mathsf{E}(X_2) = 0, \quad \mathsf{Var}(X_1) = 1, \quad \mathsf{Var}(X_2) = 1, \quad \mathsf{Corr}(X_1, X_2) = \rho.$$

Therefore,

$$\mathsf{E}(Y_1) = \mathsf{E}(X_1 + X_2 + 1) = \mathsf{E}(X_1) + \mathsf{E}(X_2) + 1 = 0 + 0 + 1 = 1,$$

$$\mathsf{E}(Y_2) = \mathsf{E}(2X_1 - X_2 - 3) = 2\mathsf{E}(X_1) - \mathsf{E}(X_2) - 3 = 2 \cdot 0 - 0 - 3 = -3,$$

$$\mathsf{Cov}(X_1, X_2) = \mathsf{Corr}(X_1, X_2)\sqrt{\mathsf{Var}(X_1)\mathsf{Var}(X_2)} = \rho\sqrt{1 \cdot 1} = \rho,$$

$$\mathsf{Var}(Y_1) = \mathsf{Var}(X_1 + X_2 + 1) = \mathsf{Var}(X_1) + \mathsf{Var}(X_2) + 2\mathsf{Cov}(X_1, X_2) = 1 + 1 + 2\rho = 2 + 2\rho,$$

$$\mathsf{Var}(Y_2) = \mathsf{Var}(2X_1 - X_2 - 3) = 4\mathsf{Var}(X_1) + \mathsf{Var}(X_2) - 4\mathsf{Cov}(X_1, X_2) = 4 + 1 - 4\rho = 5 - 4\rho,$$

$$\mathsf{Cov}(Y_1, Y_2) = \mathsf{Cov}(X_1 + X_2 + 1, 2X_1 - X_2 - 3) = \mathsf{Cov}(X_1 + X_2, 2X_1 - X_2) = 1 + \rho,$$

$$\mathsf{Corr}(Y_1, Y_2) = \frac{\mathsf{Cov}(Y_1, Y_2)}{\sqrt{\mathsf{Var}(Y_1)\mathsf{Var}(Y_2)}} = \frac{1 + \rho}{\sqrt{(2 + 2\rho)(5 - 4\rho)}},$$

Thus, $(Y_1, Y_2) \sim$ bivariate normal$(1, -3, 2 + 2\rho, 5 - 4\rho, \frac{1+\rho}{\sqrt{(2+2\rho)(5-4\rho)}})$.

Problem 5. *Statistical Inference* by Casella and Berger, 2nd Edition, Chapter 4, Exercise 32

32. (a) For the hierarchical model

$$Y|\Lambda \sim \text{Poisson}(\Lambda), \qquad \text{and} \qquad \Lambda \sim \text{Gamma}(\alpha, \beta),$$

find the marginal distribution, mean, and variance of Y. Show that the marginal distribution of Y is a negative binomial distribution if α is an integer. Also write down the joint "density" function (as described on Slide 283). Remark: the distribution you get for Y is an extension of the negative binomial defined in the notes and book, when the first parameter is not an integer. Nevertheless, it is still called "negative binomial".

(b) Determine the conditional distribution of Λ given Y. [This would be of interest in a Bayesian setting, for example.]

(a) The marginal distribution of Y is

$$\begin{aligned} f_Y(y) &= \int_0^\infty f_{Y|\Lambda}(y|\lambda) f_\Lambda(\lambda) d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda}\lambda^y}{y!} \cdot \frac{1}{\beta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}} d\lambda \\ &= \frac{1}{y!\beta^\alpha\Gamma(\alpha)} \int_0^\infty \lambda^{y+\alpha-1} e^{-\lambda(1+\frac{1}{\beta})} d\lambda \\ &= \frac{1}{y!\beta^\alpha\Gamma(\alpha)} \cdot \frac{\Gamma(y+\alpha)}{(1+\frac{1}{\beta})^{y+\alpha}} \\ &= \frac{\Gamma(y+\alpha)}{y!\Gamma(\alpha)} \cdot \left(\frac{1}{1+\beta}\right)^\alpha \cdot \left(\frac{\beta}{1+\beta}\right)^y, \quad y = 0, 1, 2, \ldots \end{aligned}$$

The mean of Y is

$$\mathsf{E}(Y) = \mathsf{E}(\mathsf{E}(Y|\Lambda)) = \mathsf{E}(\Lambda) = \alpha\beta,$$

and the variance of Y is

$$\mathsf{Var}(Y) = \mathsf{E}(\mathsf{Var}(Y|\Lambda)) + \mathsf{Var}(\mathsf{E}(Y|\Lambda)) = \mathsf{E}(\Lambda) + \mathsf{Var}(\Lambda) = \alpha\beta + \alpha\beta^2 = \alpha\beta(1+\beta).$$

If α is an integer, then the marginal distribution of Y is a negative binomial distribution with parameters $r = \alpha$ and $p = \frac{1}{1+\beta}$.

(b) The joint "density" function of (Y, Λ) is

$$f_{Y,\Lambda}(y,\lambda) = f_{Y|\Lambda}(y|\lambda) f_\Lambda(\lambda) = \frac{e^{-\lambda}\lambda^y}{y!} \cdot \frac{1}{\beta^\alpha\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}}, \quad y = 0, 1, 2, \ldots, \quad \lambda > 0.$$

The conditional distribution of Λ given Y is

$$\begin{aligned} f_{\Lambda|Y}(\lambda|y) &= \frac{f_{Y,\Lambda}(y,\lambda)}{f_Y(y)} \\ &= \frac{\frac{e^{-\lambda}\lambda^y}{y!} \cdot \frac{1}{\beta^\alpha\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}}}{\frac{\Gamma(y+\alpha)}{y!\Gamma(\alpha)} \cdot \left(\frac{1}{1+\beta}\right)^\alpha \cdot \left(\frac{\beta}{1+\beta}\right)^y} \\ &= \frac{(1+\beta)^{y+\alpha}}{\beta^{y+\alpha}\Gamma(y+\alpha)} \lambda^{y+\alpha-1} e^{-\lambda(1+\frac{1}{\beta})}, \quad \lambda > 0. \end{aligned}$$

This is the pdf of a Gamma distribution with parameters $y + \alpha$ and $\frac{\beta}{1+\beta}$.

Problem 6. (Copulas)

(a) Use the Farlie-Morgenstern copula to obtain a joint cdf with marginal cdf's $F(x) = x^4$, $0 \le x \le 1$, and $G(y) = 1 - (1-y)^2$, $0 \le y \le 1$. What is the pdf?

(b) Let $\alpha \in [-1, 1]$ and

$$f(u,v) = 1 + \alpha\cos(2\pi(u-v)), \quad \text{for } 0 \le u \le 1,\ 0 \le v \le 1,$$

Show that $f(u,v)$ is a bivariate pdf. Show that it is the pdf for a copula by finding the marginal pdfs. (Use the fact that $\cos(x)$ is periodic with period 2π.)

(a) The Farlie-Morgenstern copula is given by

$$G(u,v) = uv(1+\alpha(1-u)(1-v)), \quad \text{for } 0<u<1,\ 0<v<1.$$

Therefore, the joint cdf with marginal cdf's $F(x) = x^4$ and $G(y) = 1-(1-y)^2$ is

$$H(x,y) = G(F(x),G(y)) = x^4\left(1-(1-y)^2\right)\left[1+\alpha(1-x^4)(1-(1-(1-y)^2))\right].$$

The pdf is

$$\begin{aligned}
h(x,y) &= \frac{\partial^2 H(x,y)}{\partial x \partial y} \\
&= \frac{\partial}{\partial x}\left(4x^3(1-(1-y)^2)\left[1+\alpha(1-x^4)(1-(1-(1-y)^2))\right]\right) \\
&= 4x^3(2(1-y))\left[1+\alpha(1-x^4)(2y-y^2)\right] + x^4(2(1-y))\left[-4\alpha x^3(2y-y^2)\right] \\
&= 8x^3(1-y)\left[1+\alpha(1-x^4)(2y-y^2) - 2\alpha x^4(2y-y^2)\right] \\
&= 8x^3(1-y)\left[1+\alpha(2y-y^2)(1-3x^4)\right].
\end{aligned}$$

(b) To show that $f(u,v)$ is a bivariate pdf, we need to verify that

$$\int_0^1\int_0^1 f(u,v)\,dudv = 1.$$

We have

$$\begin{aligned}
\int_0^1\int_0^1 f(u,v)\,dudv &= \int_0^1\int_0^1 (1+\alpha\cos(2\pi(u-v)))\,dudv \\
&= \int_0^1 \left[u + \frac{\alpha}{2\pi}\sin(2\pi(u-v)) \Big|_{u=0}^{u=1} \right] dv \\
&= \int_0^1 (1+0)\,dv = 1.
\end{aligned}$$

Next, we find the marginal pdfs:

$$f_U(u) = \int_0^1 f(u,v)\,dv, \quad f_V(v) = \int_0^1 f(u,v)\,du.$$

We have

$$\begin{aligned}
f_U(u) &= \int_0^1 (1+\alpha\cos(2\pi(u-v)))\,dv \\
&= \left[v + \frac{\alpha}{2\pi}\sin(2\pi(u-v)) \Big|_{v=0}^{v=1} \right] \\
&= 1+0 = 1,
\end{aligned}$$

Similarly,

$$\begin{aligned} f_V(v) &= \int_0^1 (1 + \alpha \cos(2\pi(u - v)))du \\ &= \left[u + \frac{\alpha}{2\pi} \sin(2\pi(u - v)) \Big|_{u=0}^{u=1} \right] \\ &= 1 + 0 = 1. \end{aligned}$$

Thus, the marginal pdfs are uniform on [0, 1], confirming that $f(u, v)$ is a copula pdf.

Problem 7. Let $X_1, \ldots, X_n$ be iid gamma(α, β) random variables, with common mgf

$$M_X(t) = (1 - \beta t)^{-\alpha}.$$

(a) Recall Theorem 5.6 in the notes and find the mgf for $\bar{X}$. Identify it as the mgf for a gamma distribution. Confirm it has the right mean and variance (according to what we know generally about $\bar{X}$).

(b) Compute $E(1/\bar{X})$ and $\text{Var}(1/\bar{X})$. Hint: for a gamma rv T, $E(T^{-j})$ can be expressed in terms of a gamma integral as long as j is not too large. Use the recursion formula for $\Gamma(z)$ to simplify, when $j = 1$ or $j = 2$. [The mean and variance of $1/\bar{X}$ would be of interest, for example, when α is known and you want to estimate the rate parameter $1/\beta$ with $\alpha/\bar{X}$.]

(c) Find the pdf for $Y = 1/\bar{X}$. [This is known as an inverse gamma distribution.]

(d) Let $\hat{\mu}'_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$. Find $E(\hat{\mu}'_2)$ and $\text{Var}(\hat{\mu}'_2)$ (simplify).

(a) **Theorem 5.6.** Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with mgfs $M_1, M_2, \ldots, M_n$, respectively.
Then $X_1 + \cdots + X_n$ has mgf $M(t) = \prod_{i=1}^n M_i(t)$.
In particular, if $X_1, X_2, \ldots, X_n$ are iid then $X_1 + \cdots + X_n$ has mgf $M(t) = (M_1(t))^n$ and $\bar{X}$ has mgf $M_{\bar{X}}(t) = (M_1(t/n))^n$.

Thus, the mgf for $\bar{X}$ is

$$M_{\bar{X}}(t) = \left(M_1 \left(\frac{t}{n} \right) \right)^n = \left(1 - \beta \cdot \frac{t}{n} \right)^{-n\alpha} = \left(1 - \frac{\beta}{n} t \right)^{-n\alpha}.$$

This is the mgf of a gamma distribution with shape parameter $n\alpha$ and scale parameter $\frac{\beta}{n}$. Therefore, the mean and variance of $\bar{X}$ are

$$\mathsf{E}(\bar{X}) = n\alpha \cdot \frac{\beta}{n} = \alpha\beta, \quad \mathsf{Var}(\bar{X}) = n\alpha \cdot \left(\frac{\beta}{n} \right)^2 = \frac{\alpha\beta^2}{n}.$$

(b) We have

$$\mathsf{E}\left(\frac{1}{X}\right) = \int_0^\infty \frac{1}{x} \cdot \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} x^{n\alpha-1} e^{-\frac{n}{\beta}x} dx$$
$$= \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} \int_0^\infty x^{n\alpha-2} e^{-\frac{n}{\beta}x} dx$$
$$= \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} \cdot \frac{\Gamma(n\alpha-1)}{(n/\beta)^{n\alpha-1}}$$
$$= \frac{n}{\beta(n\alpha-1)}, \quad n\alpha > 1.$$

Similarly,

$$\mathsf{E}\left(\frac{1}{X^2}\right) = \int_0^\infty \frac{1}{x^2} \cdot \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} x^{n\alpha-1} e^{-\frac{n}{\beta}x} dx$$
$$= \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} \int_0^\infty x^{n\alpha-3} e^{-\frac{n}{\beta}x} dx$$
$$= \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} \cdot \frac{\Gamma(n\alpha-2)}{(n/\beta)^{n\alpha-2}}$$
$$= \frac{n^2}{\beta^2(n\alpha-1)(n\alpha-2)}, \quad n\alpha > 2.$$

Therefore,

$$\mathsf{Var}\left(\frac{1}{X}\right) = \mathsf{E}\left(\frac{1}{X^2}\right) - \left(\mathsf{E}\left(\frac{1}{X}\right)\right)^2 = \frac{n^2}{\beta^2(n\alpha-1)(n\alpha-2)} - \left(\frac{n}{\beta(n\alpha-1)}\right)^2.$$

(c) The pdf for $Y = 1/X$ is

$$f_Y(y) = f_X\left(\frac{1}{y}\right) \cdot \left|\frac{d}{dy}\left(\frac{1}{y}\right)\right|$$
$$= \frac{1}{\Gamma(n\alpha)(\beta/n)^{n\alpha}} \left(\frac{1}{y}\right)^{n\alpha-1} e^{-\frac{n}{\beta}\cdot\frac{1}{y}} \cdot \frac{1}{y^2}$$
$$= \frac{(n/\beta)^{n\alpha}}{\Gamma(n\alpha)} y^{-n\alpha-1} e^{-\frac{n}{\beta y}}, \quad y > 0.$$

(d) We have

$$\mathsf{E}(\hat{\mu}_2') = \mathsf{E}\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right) = \frac{1}{n}\sum_{i=1}^n \mathsf{E}(X_i^2) = \mathsf{E}(X_1^2).$$

Since $X_1 \sim \text{gamma}(\alpha, \beta)$, we have

$$\mathsf{E}(X_1^2) = \mathsf{Var}(X_1) + (\mathsf{E}(X_1))^2 = \alpha\beta^2 + (\alpha\beta)^2 = \alpha\beta^2(1+\alpha).$$

Next, we compute $\mathsf{Var}(\hat{\mu}_2')$:

$$\mathsf{Var}(\hat{\mu}_2') = \mathsf{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathsf{Var}(X_i^2) = \frac{1}{n}\mathsf{Var}(X_1^2).$$

To compute $\mathsf{Var}(X_1^2)$, we first compute $\mathsf{E}(X_1^4)$:

$$\begin{aligned}
\mathsf{E}(X_1^4) &= \int_0^\infty x^4 \cdot \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}\, dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{\alpha+3} e^{-\frac{x}{\beta}}\, dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot \Gamma(\alpha+4)\beta^{\alpha+4} \\
&= \frac{\Gamma(\alpha+4)}{\Gamma(\alpha)}\beta^4.
\end{aligned}$$

Therefore,

$$\mathsf{Var}(X_1^2) = \mathsf{E}(X_1^4) - (\mathsf{E}(X_1^2))^2 = \frac{\Gamma(\alpha+4)}{\Gamma(\alpha)}\beta^4 - (\alpha\beta^2(1+\alpha))^2.$$

Thus,

$$\mathsf{Var}(\hat{\mu}_2') = \frac{1}{n}\mathsf{Var}(X_1^2) = \frac{1}{n}\left(\frac{\Gamma(\alpha+4)}{\Gamma(\alpha)}\beta^4 - (\alpha\beta^2(1+\alpha))^2\right).$$