STAT 712 hw 7

Covariance, hierarchical models, inequalities

Do problems 4.32, 4.42, 4.43, 4.54, 4.58, 4.63 from CB.

1. For $Z_1, \dots, Z_p \sim \text{Normal}(0,1)$, not necessarily independent, prove the maximal inequality

$$\mathbb{E} \max_{1 \le j \le p} Z_j \le \sqrt{2 \log p}.$$

Use these steps:

(a) Show that for all $t \in \mathbb{R}$ we have $\exp(t \cdot \mathbb{E} \max_{1 \le j \le p} Z_j) \le pe^{t^2/2}$. *Hint: Begin with Jensen's.*

We have

$$\exp(t \cdot \mathbb{E} \max_{1 \le j \le p} Z_j) \le \mathbb{E} \exp(t \max_{1 \le j \le p} Z_j) = \mathbb{E} \max_{1 \le j \le p} \exp(tZ_j) \le \mathbb{E} \sum_{j=1}^{p} e^{tZ_j} = pe^{t^2/2},$$

where the first inequality comes from Jensen's inequality and the second from the fact that the sum cannot be less than the maximum of a set of positive numbers.

(b) Find the value of t yielding the best possible upper bound on $\mathbb{E} \max_{1 \le j \le p} Z_j$.

Taking the natural logarithm of both sides, we see that

$$t \cdot \mathbb{E} \max_{1 \le j \le p} Z_j \le \log p + \frac{t^2}{2} \quad \iff \quad \mathbb{E} \max_{1 \le j \le p} Z_j \le \frac{\log p}{t} + \frac{t}{2}.$$

The value of t which makes the right-hand side the smallest is $t = \sqrt{2 \log p}$ (take the derivative with respect to t, set it equal to 0, and solve for t; check that 2nd derivative is positive). Plugging this in for t gives the desired inequality.

2. Let $X_1 \sim \text{Normal}(m_1, \kappa^{-1})$ and $X_2 \sim \text{Normal}(m_2, \kappa^{-1})$ be independent rvs.

(a) Let the rv pair (R, Θ) be defined by $X_1 = R \cos \Theta$ and $X_2 = R \sin \Theta$, where $R > 0$ and $\Theta \in [-\pi, \pi)$. In addition, represent m_1 and m_2 as $m_1 = s \cdot \cos \mu$ and $m_2 = s \cdot \sin \mu$ for some $s > 0$ and $\mu \in [-\pi, \pi)$. Give the joint pdf of (R, Θ).

The joint pdf of (X_1, X_2) is given by

$$f(x_1, x_2) = \frac{\kappa}{2\pi} \exp\left[-\frac{\kappa}{2}(x_1 - m_1)^2 + (x_2 - m_2)^2\right]$$

The Jacobian of the transformation defined by $(x_1, x_2) = (r \cos \theta, r \sin \theta)$ is given by

$$\left| \begin{array}{cc} \frac{\partial}{\partial r} r \cos \theta & \frac{\partial}{\partial r} r \sin \theta \\ \frac{\partial}{\partial \theta} r \cos \theta & \frac{\partial}{\partial \theta} r \sin \theta \end{array} \right| = \left| \begin{array}{cc} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{array} \right| = r \cos^2 \theta + r \sin^2 \theta = r.$$

The joint pdf of (R, Θ), by the multivariate transformation method, is given by

$$\begin{aligned} f(r,\theta) &= \frac{r\kappa}{2\pi}\exp\left[-\frac{\kappa}{2}(r\cos\theta - s\cos\mu)^2 + (r\sin\theta - s\sin\mu)^2\right] \\ &= \frac{r\kappa}{2\pi}\exp\left[-\frac{\kappa}{2}(r^2 + s^2 - 2rs(\cos\theta\cos\mu + \sin\theta\sin\mu)\right] \\ &= \frac{r\kappa}{2\pi}\exp\left[-\frac{\kappa}{2}(r^2 + s^2 - 2rs\cos(\theta-\mu)\right] \end{aligned}$$

for $r > 0$ and $\theta \in [-\pi, \pi)$.

(b) Note that if $s = 1$ the point (m_1, m_2) lies on the unit circle and if $R = 1$ the point (X_1, X_2) lies on the unit circle. Show that, under $s = 1$, the conditional density of Θ given $R = 1$ is given by

$$f(\theta|R=1) = \frac{e^{\kappa\cos(\theta-\mu)}}{\int_{-\pi}^{\pi} e^{\kappa\cos(\theta'-\mu)}d\theta'}\mathbf{1}(-\pi \leq \theta < \pi).$$

This is the pdf of the von Mises distribution, which is often used for modeling directional data.

The conditional density of θ given $R = r$ is given by $f(\theta|r) = f(\theta, r)/f(r)$, where $f(r) = \int_{-\pi}^{\pi} f(\theta, r)d\theta$. Setting $s = 1$ and plugging $r = 1$ into $f(\theta, r)$, we have (abusing notation for the sake of clarity)

$$f(\theta, r=1) = \frac{\kappa}{2\pi}\exp\left[-\kappa(1-\cos(\theta-\mu)\right] = \frac{\kappa}{2\pi}e^{-\kappa}e^{\kappa\cos(\theta-\mu)}.$$

Now we write

$$f(\theta|r=1) = f(\theta,1)/f(r=1) = \frac{\frac{\kappa}{2\pi}e^{-\kappa}e^{\kappa\cos(\theta-\mu)}}{\int_{-\pi}^{\pi}\frac{\kappa}{2\pi}e^{-\kappa}e^{\kappa\cos(\theta'-\mu)}d\theta'} = \frac{e^{\kappa\cos(\theta-\mu)}}{\int_{-\pi}^{\pi} e^{\kappa\cos(\theta'-\mu)}d\theta'}.$$

3. (Optional) Additional problems from CB: 4.33, 4.40.

Problems 4.32, 4.42, 4.54, 4.58 from CB.

$\boxed{4.32}$

(a) Let $Y|\Lambda \sim \text{Poisson}(\Lambda)$

$\Lambda \sim \text{Gamma}(\alpha, \beta)$

For $y = 0, 1, 2, \ldots$ the marginal pmf of Y is given by

$$p_Y(y) = \int_0^\infty \frac{e^{-\lambda}\lambda^y}{y!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} \, d\lambda$$

$$= \frac{\left(\frac{\beta}{\beta+1}\right)^{y+\alpha} \Gamma(y+\alpha)}{\Gamma(\alpha)\beta^\alpha \, y!} \underbrace{\int_0^\infty \frac{1}{\Gamma(y+\alpha)\left(\frac{\beta}{\beta+1}\right)^{y+\alpha}} \lambda^{y+\alpha-1} e^{-\lambda/\left(\frac{\beta}{\beta+1}\right)} \, d\lambda}_{=1,\ \text{integral over Gamma}\left(y+\alpha,\ \frac{\beta}{\beta+1}\right)\ \text{pdf}}$$

$$= \frac{\Gamma(y+\alpha)}{\Gamma(\alpha)\, y!} \left(\frac{1}{\beta+1}\right)^\alpha \left(\frac{\beta}{\beta+1}\right)^y$$

If α is an integer, we have the negative binomial pmf

$$p_Y(y) = \binom{\alpha+y-1}{y} \left(\frac{1}{\beta+1}\right)^\alpha \left(1 - \frac{1}{\beta+1}\right)^y ,$$

which is the pmf of Y if $Y = \#$ failures before the α^{th} success, when successes occur with probability $\frac{1}{\beta+1}$.

Also

$$\mathbb{E}\,Y = \mathbb{E}\left(\mathbb{E}[Y|\Lambda]\right) = \mathbb{E}(\Lambda) = \alpha\beta$$

$$\text{Var}\,Y = \text{Var}\left(\mathbb{E}[Y|\Lambda]\right) + \mathbb{E}\left(\text{Var}[Y|\Lambda]\right)$$

$$= \operatorname{Var}(\Lambda) + \mathbb{E}(\Lambda)$$

$$= \alpha\beta^2 + \alpha\beta$$

$$= \alpha\beta(\beta+1).$$

(b) Let

$$Y \mid N \sim \text{Binomial}(N, p)$$

$$N \mid \Lambda \sim \text{Poisson}(\Lambda)$$

$$\Lambda \sim \text{Gamma}(\alpha, \beta)$$

Then N has marginal pmf

$$p_N(n) = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\, n!} \left(\frac{1}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{n} ,$$

and the marginal of Y is given by

$$p_Y(y) = \sum_{n=0}^{\infty} \binom{n}{y} p^y (1-p)^{n-y} \mathbb{1}(y \in \{0,1,\dots,n\}) \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\, n!} \left(\frac{1}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{n}$$

$$= \sum_{n=y}^{\infty} \binom{n}{y} p^y (1-p)^{n-y} \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\, n!} \left(\frac{1}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{n}$$

$m = n-y$
$n = m+y$

$$= \sum_{m=0}^{\infty} \binom{m+y}{y} p^y (1-p)^{m} \frac{\Gamma(m+y+\alpha)}{\Gamma(\alpha)\,(m+y)!} \left(\frac{1}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{m+y}$$

$$= \frac{\frac{\Gamma(y+\alpha)}{y!\,\Gamma(\alpha)} \left(\frac{p\beta}{\beta+1}\right)^{y} \left(\frac{1}{\beta+1}\right)^{\alpha}}{\Big(\underbrace{1 - \frac{\beta(1-p)}{\beta+1}}_{= \frac{1+p\beta}{\beta+1}}\Big)^{y+\alpha}} \underbrace{\sum_{m=0}^{\infty} \frac{\Gamma(m+y+\alpha)}{\Gamma(y+\alpha)\, m!} \left(1 - \frac{\beta(1-p)}{\beta+1}\right)^{y+\alpha} \left(\frac{\beta(1-p)}{\beta+1}\right)^{m}}_{=1,\ \text{sum over neg. binom pmf}}$$

$$= \frac{\Gamma(y+\alpha)}{\Gamma(\alpha)\, y!} \left(\frac{p\beta}{1+p\beta}\right)^{y} \left(\frac{1}{1+p\beta}\right)^{\alpha}.$$

When α is an integer the pmf becomes

$$p_Y(y) = \binom{y+\alpha-1}{y} \left(\frac{1}{1+p\beta}\right)^{\alpha} \left(1 - \frac{1}{1+p\beta}\right)^{y},$$

which is the dist of $Y = \#$ failures before α^{th} success under success probability $1/(1+p\beta)$.

$\boxed{4.42}$ Let $X \perp\!\!\!\perp Y$, $\mathbb{E}X = \mu_x$, $\mathbb{E}Y = \mu_y$, $\operatorname{Var} X = \sigma_x^2$, $\operatorname{Var} Y = \sigma_y^2$.

We have

$$\operatorname{Corr}(XY, Y) = \frac{\operatorname{Cov}(XY, Y)}{\sqrt{\operatorname{Var}(XY)\operatorname{Var} Y}} .$$

where

$$\begin{aligned}
\operatorname{Cov}(XY, Y) &= \mathbb{E}(XY \cdot Y) - \mathbb{E}XY \cdot \mathbb{E}Y \\
&= \mathbb{E}X\, \mathbb{E}Y^2 - \mathbb{E}X \cdot (\mathbb{E}Y)^2 \\
&= \mu_x \sigma_y^2
\end{aligned}$$

$$\begin{aligned}
\operatorname{Var}(XY) &= \mathbb{E}[(XY)^2] - [\mathbb{E}(XY)]^2 \\
&= \mathbb{E}X^2 \cdot \mathbb{E}Y^2 - (\mathbb{E}X)^2(\mathbb{E}Y)^2 \\
&= (\sigma_x^2 + \mu_x^2)(\sigma_y^2 + \mu_y^2) - \mu_x^2 \mu_y^2 \\
&= \sigma_x^2 \sigma_y^2 + \sigma_x^2 \mu_y^2 + \sigma_y^2 \mu_x^2
\end{aligned}$$

Plugging in these expressions gives

$$\begin{aligned}
\operatorname{Corr}(XY, Y) &= \frac{\mu_x \sigma_y^2}{\sqrt{(\sigma_x^2 \sigma_y^2 + \sigma_x^2 \mu_y^2 + \sigma_y^2 \mu_x^2)\, \sigma_y^2}} \\
&= \frac{\mu_x \sigma_y}{\sqrt{(\sigma_x^2 \sigma_y^2 + \sigma_x^2 \mu_y^2 + \sigma_y^2 \mu_x^2)}}
\end{aligned}$$

$\boxed{4.54}$ Let $X_1, \ldots, X_n \overset{ind}{\sim} \text{Unif}(0,1)$. Find pdf of $Y = \prod_{i=1}^{n} X_i$.

For $y \in (0,1)$, the cdf of Y is given by

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P\left(\prod_{i=1}^{n} X_i \le y \right) \\
&= P\left(\log \prod_{i=1}^{n} X_i \le \log y \right) \\
&= P\left(\sum_{i=1}^{n} \log X_i \le \log y \right) \\
&= P\left(\sum_{i=1}^{n} (-\log X_i) \geq -\log y \right) \\
&= P\left(\sum_{i=1}^{n} W_i \geq -\log y \right), \quad W_1, \ldots, W_n \overset{ind}{\sim} \text{Exponential}(1) \\
&= P\left(G \geq -\log y \right), \quad \text{where } G \sim \text{Gamma}(n,1) \\
&= 1 - P\left(G < -\log y \right) \\
&= 1 - F_G(-\log y), \quad \text{where } F_G \text{ the Gamma}(n,1) \text{ cdf.}
\end{aligned}
$$

Now, for $y \in (0,1)$, we have

$$
\begin{aligned}
f_Y(y) &= \frac{d}{dy} F_G(-\log y) \\
&= f_G(-\log y)\left(-\frac{1}{y}\right), \quad f_G \text{ the Gamma}(n,1) \text{ pdf} \\
&= \frac{1}{\Gamma(n)} (-\log y)^{n-1} e^{-(-\log y)} \frac{1}{y} \\
&= \frac{1}{(n-1)!} (-\log y)^{n-1}.
\end{aligned}
$$

$\boxed{4.58}$ Let X and Y have finite variances. Show

(a) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(X, \mathbb{E}[Y|X])$. We have

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y \\ &= \mathbb{E}\big(\mathbb{E}[XY|X]\big) - \mathbb{E}X\,\mathbb{E}\big(\mathbb{E}[Y|X]\big) \\ &= \mathbb{E}\big(X\cdot\mathbb{E}[Y|X]\big) - \mathbb{E}X\cdot\mathbb{E}\big(\mathbb{E}[Y|X]\big) \\ &= \mathrm{Cov}\big(X, \mathbb{E}[Y|X]\big).\end{aligned}$$

(b) X and $Y - \mathbb{E}[Y|X]$ are uncorrelated. We have

$$\begin{aligned}\mathrm{Cov}\big(X, Y-\mathbb{E}[Y|X]\big) &= \mathbb{E}\big[X\cdot(Y-\mathbb{E}[Y|X])\big] - \mathbb{E}X\cdot\mathbb{E}\big(Y-\mathbb{E}[Y|X]\big) \\ &= \mathbb{E}XY - \mathbb{E}\big(X\cdot\mathbb{E}[Y|X]\big) - \mathbb{E}X\cdot\mathbb{E}Y + \mathbb{E}X\cdot\mathbb{E}\big(\mathbb{E}[Y|X]\big) \\ &= \mathrm{Cov}(X,Y) - \underbrace{\mathrm{Cov}\big(X,\mathbb{E}[Y|X]\big)}_{=\mathrm{Cov}(X,Y)\text{ from (a)}} \\ &= 0.\end{aligned}$$

(c) $\mathrm{Var}\big(Y - \mathbb{E}[Y|X]\big) = \mathbb{E}\big(\mathrm{Var}[Y|X]\big)$. We have

$$\begin{aligned}\mathrm{Var}\big(Y-\mathbb{E}[Y|X]\big) &= \mathbb{E}\Big(Y-\mathbb{E}[Y|X] - \underbrace{\mathbb{E}\big(Y-\mathbb{E}[Y|X]\big)}_{=0}\Big)^2 \\ &= \mathbb{E}\big(Y-\mathbb{E}[Y|X]\big)^2 \\ &= \mathbb{E}\Big(\mathbb{E}\big[(Y-\mathbb{E}[Y|X])^2\,|\,X\big]\Big) \\ &= \mathbb{E}\big(\mathrm{Var}[Y|X]\big).\end{aligned}$$