Statistics 351–Probability I
Fall 2007 (200730)
Final Exam Solutions

Instructor: Michael Kozdron

1. (a) We see that $f_{X,Y}(x,y) \ge 0$ for all x, y, and that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = \int_0^1\int_0^y 8xy\,dx\,dy = \int_0^1 4y^3\,dy = y^4\Big|_0^1 = 1.$$

Thus, $f_{X,Y}$ is a legitimate density.

1. (b) We compute

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_x^1 8xy\,dy = 4x(1-x^2), \quad 0 < x < 1.$$

1. (c) We compute

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^1 x\cdot 4x(1-x^2)\,dx = \frac{4}{3} - \frac{4}{5} = \frac{8}{15}.$$

1. (d) We compute

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{8xy}{4x(1-x^2)} = \frac{2y}{(1-x^2)}, \quad x < y < 1.$$

1. (e) We compute

$$E(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X=x}(y)\,dy = \int_x^1 y\cdot\frac{2y}{(1-x^2)}\,dy = \frac{2(1-x^3)}{3(1-x^2)}.$$

1. (f) Using properties of conditional expectation (Theorem II.2.1), we compute

$$E(Y) = E(\,E(Y|X)\,) = E\left(\frac{2(1-X^3)}{3(1-X^2)}\right) = \int_0^1 \frac{2(1-x^3)}{3(1-x^2)}\cdot 4x(1-x^2)dx = \frac{8}{3}\int_0^1 x - x^4 dx = \frac{4}{5}.$$

2. (a) Let

$$B = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

so that $\mathbf{Y} = B\mathbf{X}$. By Theorem V.3.1, $\mathbf{Y}$ is MVN with mean

$$B\boldsymbol{\mu} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and covariance matrix

$$B\boldsymbol{\Lambda}B' = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 14 & 8 \\ 8 & 5 \end{pmatrix}.$$

2. (b) Note that

$$\det\begin{pmatrix}14 & 8\\ 8 & 5\end{pmatrix} = 70 - 64 = 6$$

so that

$$\begin{pmatrix}14 & 8\\ 8 & 5\end{pmatrix}^{-1} = \begin{pmatrix}\frac{5}{6} & -\frac{8}{6}\\ -\frac{8}{6} & \frac{14}{6}\end{pmatrix}.$$

Thus, we can conclude

$$f_{Y_1,Y_2}(y_1,y_2) = \frac{1}{2\pi}\cdot\frac{1}{\sqrt{6}}\exp\left\{-\frac{1}{2}\left(\frac{5}{6}y_1^2 - \frac{8}{3}y_1y_2 + \frac{7}{3}y_2^2\right)\right\}.$$

2. (c) Since

$$\mathbf{\Lambda} = \begin{pmatrix}14 & 8\\ 8 & 5\end{pmatrix}$$

we can immediately conclude that

$$\varphi(t_1,t_2) = \exp\left\{-\frac{1}{2}\left(14t_1^2 + 16t_1t_2 + 5t_2^2\right)\right\}.$$

3. (a) Using the results of Section V.6 (in particular, equation (6.2) on page 130) we know that

$$X_2|X_1 = x \in \mathcal{N}\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2)\right).$$

Since $\sigma_1^2 = 1$, $\sigma_2^2 = 25$, we conclude that

$$\rho = \frac{\mathrm{Cov}(X_1,X_2)}{\sigma_1\sigma_2} = \frac{\alpha}{5}.$$

Therefore,

$$16 = \mathrm{Var}(X_2|X_1) = \sigma_2^2(1-\rho^2) = 25\left(1-\frac{\alpha^2}{25}\right) = 25 - \alpha^2$$

implying that $\alpha^2 = 9$. Hence, the two possible values of α are $\alpha = 3$ and $\alpha = -3$.

3. (b) From **(a)**, we conclude that

$$1 = \mathbb{E}(X_2|X_1 = 6) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1) = \beta + \frac{\alpha}{5}\cdot\frac{5}{1}(6-5) = \beta + \alpha.$$

Therefore, if $\alpha = 3$, then $\beta = -2$ and if $\alpha = -3$, then $\beta = 4$.

4. (a) In order to find the eigenvalues of $\mathbf{\Lambda}$, we must find those values of λ such that $\det(\mathbf{\Lambda} - \lambda I) = 0$. Therefore,

$$\det(\mathbf{\Lambda}-\lambda I) = \det\begin{pmatrix}6-\lambda & 2\\ 2 & 9-\lambda\end{pmatrix} = (6-\lambda)(9-\lambda)-4 = \lambda^2 - 15\lambda + 54 - 4 = \lambda^2 - 15\lambda + 50 = (\lambda-5)(\lambda-10)$$

so that the eigenvalues of $\mathbf{\Lambda}$ are $\lambda_1 = 5$ and $\lambda_2 = 10$.

4. (b) Since $\lambda_1 = 5$,

$$(\mathbf{\Lambda} - \lambda_1 I \,|\, 0) = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

and since $\lambda_2 = 10$,

$$(\mathbf{\Lambda} - \lambda_2 I \,|\, 0) = \left(\begin{array}{cc|c} -4 & 2 & 0 \\ 2 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

we conclude that eigenvectors for λ_1 and λ_2 are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

respectively. Therefore, the diagonal matrix is

$$D = \operatorname{diag}(\lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix}$$

and the orthogonal matrix is

$$C = \left(\frac{\mathbf{v}_1}{||\mathbf{v}_1||} \; \frac{\mathbf{v}_2}{||\mathbf{v}_2||} \right) = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

since $||\mathbf{v}_1|| = ||\mathbf{v}_2|| = \sqrt{5}$.

4. (c) If $\mathbf{Y} = C'\mathbf{X}$, then by Theorem V.3.1, $\mathbf{Y}$ is MVN with mean $C'\boldsymbol{\mu}$ and covariance matrix $C'\mathbf{\Lambda}C'' = C'\mathbf{\Lambda}C = D$ using our result from **(b)**. Hence, we conclude

$$\mathbf{Y} \in N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \right).$$

4. (d) Since $\mathbf{Y}$ is multivariate normal we know from Definition I that Y_1 and Y_2 are each one-dimensional normals. We also know from Theorem V.7.1 that the components of $\mathbf{Y}$ are independent if and only if they are uncorrelated. From **(c)** we know that $\operatorname{Cov}(Y_1, Y_2) = 0$ so that Y_1 and Y_2 are, in fact, independent.

5. Notice that the density function of Y is non-zero only for $0 < y < 1$ which implies that the density function for X is non-zero only for $0 < x < 1$. Therefore, suppose that $0 < y < 1$ is fixed so that $f_{X|Y=y}(x) = 1/y$, $0 < x < y$. If we now fix $0 < x < 1$, then the range of allowable y is $x < y < 1$. Hence, by definition,

$$f_{X,Y}(x,y) = f_{X|Y=y}(x) f_Y(y) = \frac{1}{y} \cdot 20y^3(1-y) = 20y^2(1-y)$$

provided that $0 < x < y < 1$. Thus, the marginal density function of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_x^1 20y^2(1-y)dy = \left(\frac{20}{3}y^3 - \frac{20}{4}y^4 \right)\Big|_x^1 = \frac{5}{3} - \frac{20}{3}x^3 + 5x^4$$

provided that $0 < x < 1$.

6. If $U = g(X)$ and $V = h(Y)$, then solving for X and Y gives $X = g^{-1}(U)$ and $Y = h^{-1}(V)$, so that the Jacobian of this transformation is

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial}{\partial u} g^{-1}(u) & 0 \\ 0 & \frac{\partial}{\partial v} h^{-1}(v) \end{vmatrix} = \frac{\partial}{\partial u} g^{-1}(u) \cdot \frac{\partial}{\partial v} h^{-1}(v).$$

By Theorem I.2.1, the joint density of $(U, V)'$ is therefore given by

$$f_{U,V}(u, v) = f_{X,Y}(g^{-1}(u), h^{-1}(v)) \cdot |J| = f_X(g^{-1}(u)) \cdot f_Y(h^{-1}(v)) \cdot \frac{\partial}{\partial u} g^{-1}(u) \cdot \frac{\partial}{\partial v} h^{-1}(v)$$

by the assumed independence of X and Y. Since we can write $f_{U,V}(u, v)$ as a function of u only multiplied by a function of v only we conclude that U and V are, in fact, independent with

$$f_U(u) = f_X(g^{-1}(u)) \cdot \frac{\partial}{\partial u} g^{-1}(u) \quad \text{and} \quad f_V(v) = f_Y(h^{-1}(v)) \cdot \frac{\partial}{\partial v} h^{-1}(v).$$

It is worth noting that these calculations are allowed since g and h are strictly increasing and differentiable.

7. Observe that the expression

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2) \log \left(f_{\mathbf{X}}(x_1, x_2)\right) \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

exactly equals $\mathbb{E}(\log (f_{\mathbf{X}}(X_1, X_2)))$. Since

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi} \frac{1}{\sqrt{\det[\mathbf{\Lambda}]}} \exp\{-\frac{1}{2} \overline{x}' \mathbf{\Lambda}^{-1} \overline{x}\},$$

we see that

$$\mathbb{E}(\log (f_{\mathbf{X}}(X_1, X_2))) = -\log(2\pi) - \frac{1}{2} \log(\det[\mathbf{\Lambda}]) - \frac{1}{2} \mathbb{E}(\mathbf{X}' \mathbf{\Lambda}^{-1} \mathbf{X}).$$

Now, observe that if

$$\mathbf{\Lambda} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \quad \text{then} \quad \mathbf{\Lambda}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix}$$

and so

$$\mathbf{X}' \mathbf{\Lambda}^{-1} \mathbf{X} = \frac{1}{1-\rho^2} \left(\frac{X_1^2}{\sigma_1^2} - 2\rho \frac{X_1 X_2}{\sigma_1 \sigma_2} + \frac{X_2^2}{\sigma_2^2} \right).$$

Taking expected values gives

$$\mathbb{E}(\mathbf{X}' \mathbf{\Lambda}^{-1} \mathbf{X}) = \frac{1}{1-\rho^2} \mathbb{E} \left(\frac{X_1^2}{\sigma_1^2} - 2\rho \frac{X_1 X_2}{\sigma_1 \sigma_2} + \frac{X_2^2}{\sigma_2^2} \right) = \frac{1}{1-\rho^2} (1 - 2\rho^2 + 1) = 2.$$

Combining everything, we conclude that

$$\mathbb{E}(\log (f_{\mathbf{X}}(X_1, X_2))) = -\log(2\pi) - \frac{1}{2} \log(\det[\mathbf{\Lambda}]) - \frac{1}{2} \mathbb{E}(\mathbf{X}' \mathbf{\Lambda}^{-1} \mathbf{X}) = -\log(2\pi) - \frac{1}{2} \log(\det[\mathbf{\Lambda}]) - 1$$

and so

$$-\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2) \log \left(f_{\mathbf{X}}(x_1, x_2)\right) \, \mathrm{d}x_1 \, \mathrm{d}x_2 = 1 + \log(2\pi) + \frac{1}{2} \log(\det[\mathbf{\Lambda}])$$

as required.

8. If

$$X_j = \sum_{n=1}^{j} S_{n-1}(S_n - S_{n-1}).$$

then

$$X_{j+1} = X_j + S_j(S_{j+1} - S_j).$$

Therefore,

$$E(X_{j+1}|S_j) = \mathbb{E}(X_j + S_j(S_{j+1} - S_j)|S_j) = \mathbb{E}(X_j|S_j) + \mathbb{E}(S_j(S_{j+1} - S_j)|S_j) = X_j + S_j\mathbb{E}(S_{j+1}|S_j) - S_j^2$$

where we have "taken out what is known" three times. Furthermore,

$$\mathbb{E}(S_{j+1}|S_j) = \mathbb{E}(S_j + Y_{j+1}|S_j) = S_j + \mathbb{E}(Y_{j+1}) = S_j$$

where we have again "taken out what is known," and have used the facts that Y_{j+1} and S_j are independent and $\mathbb{E}(Y_{j+1}) = 0$. Combining everything gives

$$E(X_{j+1}|S_j) = X_j + S_j\mathbb{E}(S_{j+1}|S_j) - S_j^2 = X_j + S_j^2 - S_j^2 = X_j$$

which proves that $\{X_j, j = 0, 1, 2, \ldots\}$ is, in fact, a martingale.

9. The joint density of $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$ is

$$f_{X_{(1)},X_{(2)},X_{(3)},X_{(4)}}(y_1, y_2, y_3, y_4) = 4! = 24$$

provided that $0 < y_1 < y_2 < y_3 < y_4 < 1$. Thus, the joint density of $X_{(2)}, X_{(3)}$ is

$$f_{X_{(2)},X_{(3)}}(y_2, y_3) = \int_{y_3}^{1} \int_{0}^{y_2} 24 dy_1 dy_2 = 24y_2(1 - y_3)$$

provided that $0 < y_2 < y_3 < 1$. If $U = X_{(2)}/X_{(3)}$ and $V = X_{(3)}$, then solving for $X_{(2)}$ and $X_{(3)}$ gives $X_{(2)} = UV$ and $X_{(3)} = V$ so that the Jacobian of this transformation is

$$J = \begin{vmatrix} \frac{\partial y_2}{\partial u} & \frac{\partial y_2}{\partial v} \\ \frac{\partial y_3}{\partial u} & \frac{\partial y_3}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v.$$

By Theorem I.2.1, the joint density of $(U, V)'$ is therefore given by

$$f_{U,V}(u, v) = f_{X_{(2)},X_{(3)}}(uv, v) \cdot |J| = 24uv(1 - v) \cdot v = 24uv^2(1 - v)$$

provided that $0 < u < 1$ and $0 < v < 1$. Notice that U and V are, in fact, independent with $f_U(u) = 2u$, $0 < u < 1$, and $f_V(v) = 12v^2(1 - v)$, $0 < v < 1$. Finally, we see that the density of $W = U^2$ is

$$f_W(w) = \frac{d}{dw}P(W \leq w) = \frac{d}{dw}P(U \leq \sqrt{w}) = \frac{1}{2\sqrt{w}} f_U(\sqrt{w}) = \frac{1}{2\sqrt{w}} \cdot 2\sqrt{w} = 1$$

provided that $0 < w < 1$. In other words, $(X_{(2)}/X_{(3)})^2 \in U(0, 1)$ as required.

10. (a) Since $X_5 \in \text{Po}(5)$, we find

$$P(X_5 = j) = \frac{5^j}{j!} e^{-5}, \quad j = 1, 2.$$

10. (b) By adding and subtracting X_2, we compute

$$\begin{aligned}\text{Var}(X_5|X_2 = 1) = \text{Var}(X_5 - X_2 + X_2|X_2 = 1) &= \text{Var}(X_5 - X_2|X_2 = 1) + \text{Var}(X_2|X_2 = 1) \\ &= \text{Var}(X_5 - X_2) = 3\end{aligned}$$

using the fact that $X_5 - X_2 \in \text{Po}(3)$ and X_2 are independent.

10. (c) By adding and subtracting X_2, we compute

$$\text{Cov}(X_2, X_4) = \text{Cov}(X_2, X_4 - X_2 + X_2) = \text{Cov}(X_2, X_4 - X_2) + \text{Cov}(X_2, X_2) = 0 + \text{Var}(X_2)$$

using the fact that the increments $X_4 - X_2$ and X_2 are independent. Since $X_2 \in \text{Po}(2)$ we know $\text{Var}(X_2) = 2$ so that

$$\text{Cov}(X_2, X_4) = \text{Var}(X_2) = 2.$$

10. (d) By adding and subtracting X_2, we compute

$$E(X_4|X_2 = j) = E(X_4 - X_2 + X_2|X_2 = j) = E(X_4 - X_2|X_2 = j) + E(X_2|X_2 = j) = E(X_4 - X_2) + j$$

where we have used the facts that $E(X_4 - X_2|X_2 = j) = E(X_4 - X_2)$ since $X_4 - X_2$ and X_2 are independent, and $E(X_2|X_2 = j) = j$ by "taking out what is known." (See Theorems II.2.1 and II.2.2.) Since $X_4 - X_2 \in \text{Po}(2)$ we know $E(X_4 - X_2) = 2$ so that

$$E(X_4|X_2 = j) = 2 + j, \quad j = 0, 1, 2, \ldots.$$

11. (a) Let $\{X_t, t \geq 0\}$ denote the Poisson process with intensity 1 according to which Jessica buys pairs of shoes (where t measures weeks). The random number of shoes that Jessica buys in a year is X_{52}. Since $X_t \in \text{Po}(t)$ for all t by the Poisson process assumption, we conclude that $\mathbb{E}(X_{52}) = 52$.

11. (b) If we use the fact that a Poisson process resets at fixed times, then the probability that she bought 3 pairs of shoes during the first week of February given that she bought 8 pairs during the four weeks of February is

$$P(X_1 = 3|X_4 = 8) = \frac{P(X_1 = 3, X_4 = 8)}{P(X_4 = 8)} = \frac{P(X_1 = 3, X_4 - X_1 = 5)}{P(X_4 = 8)} = \frac{P(X_1 = 3)P(X_4 - X_1 = 5)}{P(X_4 = 8)}.$$

Since $X_1 \in \text{Po}(1)$, we find

$$P(X_1 = 3) = \frac{1}{3!} e^{-1},$$

since $X_4 - X_1 \in \text{Po}(3)$, we find

$$P(X_4 - X_1 = 5) = \frac{3^5}{5!} e^{-3},$$

and since $X_4 \in \text{Po}(4)$, we find

$$P(X_4 = 8) = \frac{4^8}{8!} e^{-2}.$$

Thus, the required probability is

$$P(X_1 = 3|X_4 = 8) = \frac{\frac{1}{3!} e^{-1} \cdot \frac{3^5}{5!} e^{-3}}{\frac{4^8}{8!} e^{-2}} = \frac{8!}{3!5!} \cdot \frac{3^5}{4^8}.$$