Andrew Iannaccone
Econ 203A - Matzkin
Due 10/25/07

PROB/STATS – HW #3

2.1.3 For $X, Y : S \to \mathbb{R}$, define the preimage sets

$$A = \{s \in S : X(s) \leq a\} \quad B = \{s \in S : X(s) \leq b\}$$
$$C = \{s \in S : Y(s) \leq c\} \quad D = \{s \in S : Y(s) \leq d\}$$

Note that $A \subset B$, $C \subset D$, and

$$P(a < X \leq b, c < Y \leq d) = P(B \cap A^c \cap D \cap C^c)$$

Now since in general, $P(S \cap T^c) = P(S) - P(S \cap T)$, we can rework the expression above.

$$\begin{aligned}
P &= P(B \cap A^c \cap D \cap C^c) \\
&= P(B \cap D \cap A^c) - P(B \cap D \cap A^c \cap C) \\
&= P(B \cap D) - P(B \cap D \cap A) \\
&\quad -[P(B \cap D \cap C) - P(B \cap D \cap C \cap A)] \\
&= P(B \cap D) - P(A \cap D) - P(B \cap C) + P(A \cap C) \\
&= F(b,d) - F(a,d) - F(b,c) + F(a,c)
\end{aligned}$$

2.1.4 For $(a,b,c,d) = (.5, 1, 0, 1)$, we can apply (2.1.3) to get

$$\begin{aligned}
P(.5 < X \leq 1, 0 < Y \leq 1) &= F(1,1) - F(1,0) - F(.5,1) + F(.5,0) \\
&= 1 - 1 - 1 + 0 = -1 < 0 \rightarrow\leftarrow
\end{aligned}$$

By contradiction, $F(x,y)$ cannot be a cdf of two random variables.

2.1.7 For any $z \in [0,1]$,

$$\begin{aligned}
F(z) &= \int_{\{x,y:xy\leq z\}} f(x,y) \\
&= \int_0^1 \int_0^{\min\{1,z/y\}} f(x,y)dxdy \\
&= \int_0^z \int_0^1 f(x,y)dxdy + \int_z^1 \int_0^{z/y} f(x,y)dxdy \\
&= \int_0^z 1\,dy + \int_z^1 \frac{z}{y}dy \\
&= z + \Big[z \ln y\Big]_z^1 \\
&= z - zlnz \\
\Rightarrow f(z) &= F'(z) = 1 - lnz - 1 = -lnz
\end{aligned}$$

2.2.4 Define $g(y_1, y_2) : \mathbb{R}^2 \to [0,1]$ as the joint pdf of (Y_1, Y_2), and $\mathcal{S}, \mathcal{T}$ as the supports of (X_1, X_2) and (Y_1, Y_2) respectively. Additionally define a bijective transformation $\omega(y_1, y_2) : \mathcal{T} \to \mathcal{S}$ by

$$\omega\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 y_2 \\ y_2 \end{pmatrix}$$

The Jacobian of the transformation is

$$J_\omega = \begin{pmatrix} y_2 & y_1 \\ 0 & 1 \end{pmatrix}$$

so the pdf of (Y_1, Y_2) is

$$\begin{aligned}
g(y) &= h(\omega(y))|J_\omega| \\
&= 8\omega_1(y)\omega_2(y)[y_2 \cdot 1 - y_1 \cdot 0] \\
&= 8y_1 y_2^3
\end{aligned}$$

for all $y \in \mathcal{T}$, and $g(y) = 0$ for $y \notin \mathcal{T}$. Finally, we find $\mathcal{T}$ as

$$\begin{aligned}
\mathcal{T} &= \{y \in \mathbb{R}^2 : g(y) > 0\} \\
&= \{y \in \mathbb{R}^2 : h(\omega(y)) > 0\} \\
&= \{y \in \mathbb{R}^2 : 0 < \omega_1(y) < \omega_2(y) < 1\} \\
&= \{y \in \mathbb{R}^2 : 0 < y_1 y_2 < y_2 < 1\} \\
&= (0,1) \times (0,1)
\end{aligned}$$

2.2.5 (a) Define the bijective transformation $\omega(y_1, y_2) : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\omega \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 - y_2 \\ y_2 \end{pmatrix}$$

Then $\omega(Y_1, Y_2) = (X_1, X_2)$, so the pdf of (Y_1, Y_2) is given by

$$\begin{aligned} f_{Y_1,Y_2}(y) &= f_{X_1,X_2}(\omega(y))|J_\omega| \\ &= f_{X_1,X_2}(y_1 - y_2, y_2) \end{aligned}$$

(b) The *convolution fomula* follows by integrating the joint pdf

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1, y_2)\, dy_2 \\ &= \int_{-\infty}^{\infty} f_{X_1,X_2}(y1 - y2, y2)\, dy_2 \end{aligned}$$

2.3.11 (a) Assume that $f_1(x_1) = 1$ and $f_{2|1}(x_2|x_1) = 1/x_1$ for all $0 < x_2 < x_1 < 1$.

(b) If $X_1 + X_2 \geq 1$ and $0 < X_2 < X_1 < 1$, then $X_1 > 1/2$ and $1 - X_1 \leq X_2 < X_1$. This gives us the limits of integration, so

$$\begin{aligned} P(X_1 + X_2 \geq 1) &= \int_{1/2}^{1} \int_{1-x_1}^{x_1} f_{1,2}(x_1, x_2)\, dx_2 dx_1 \\ &= \int_{1/2}^{1} \int_{1-x_1}^{x_1} f_{2|1}(x_1, x_2) f_1(x_1)\, dx_2 dx_1 \\ &= \int_{1/2}^{1} \int_{1-x_1}^{x_1} \frac{1}{x_1}\, dx_2 dx_1 \\ &= \int_{1/2}^{1} \frac{2x_1 - 1}{x_1}\, dx_1 \\ &= \Big[2x_1 - \ln x_1\Big]_{1/2}^{1} \\ &= [2 - \ln 1] - [1 - \ln \frac{1}{2}] \\ &= 1 - \ln 2 \end{aligned}$$

(c) First note the pdf for X_2 is

$$\begin{aligned}
f_2(x_2) &= \int_0^1 f_{1,2}(x_1,x_2)dx_1 \\
&= \int_0^{x_2} 0\,dx_1 + \int_{x_2}^1 \frac{1}{x_1}dx_1 \\
&= -\ln x_2 \\
\Rightarrow E(X_1|x_2) &= \int_0^1 x_1 f_{1|2}(x_1|x_2)dx_1 \\
&= \int_0^1 x_1 \frac{f_{1,2}(x_1,x_2)}{f(x_2)}dx_1 \\
&= \int_{x_2}^1 x_1 \frac{1/x_1}{-\ln x_2}dx_1 \\
&= \frac{1-x_2}{-\ln x_2}
\end{aligned}$$

2.4.11 Define $Z = (X_1 - \mu_1) + (X_2 - \mu_2)$, and note that $\mu_Z = 0$.

$$\begin{aligned}
\sigma_Z^2 &= E[(Z-\mu_Z)^2] \\
&= E[(X_1-\mu_1)^2] + 2E[(X_1-\mu_1)(X_2-\mu_2)] + E[(X_2-\mu_2)^2] \\
&= \sigma_1^2 + \sigma_2^2 + 2\sigma_{12} \\
&= 2(1+\rho)\sigma^2
\end{aligned}$$

By Chebyshev's Inequality, for all $\lambda > 0$,

$$P(Z \geq \lambda\sigma_Z) \leq \frac{1}{\lambda^2}$$

$$\Rightarrow P(Z \geq \lambda\sigma\sqrt{2(1+\rho)}) \leq \frac{1}{\lambda^2}$$

Substituting Z's definition and $k = \lambda\sqrt{2(1+\rho)}$ yields the desired expression.

2.5.7 No. Let $f_1(x_1), f_1(x_2)$ be the marginal pdfs. If X_1 and X_2 are independent, then $f_1(x_1)f_2(x_2) = f(x_1,x_2) = 1/\pi$ for all x in the support $\mathcal{S}$. Note that

$$\mathcal{S} = \{x \in \mathbb{R}^2 : (x-p)^2 < 1\} \quad \text{where } p = (1,-2)$$

Consider the nearby points $p + (\sqrt{2}, 0)$ and $p + (0, \sqrt{2})$. Both of these points are in $\mathcal{S}$, so $f_{1,2}$ is positive at each. It follows that the marginal pdfs are also positive, specifically,

$$f_1(p_1 + \sqrt{2}) > 0 \quad \text{and} \quad f_2(p_2 + \sqrt{2}) > 0$$

But this implies

$$f_{12}(p + (\sqrt{2}, \sqrt{2})) > 0,$$

which is impossible, since it lies outside $\mathcal{S}$. $\rightarrow\leftarrow$ By contradiction, X_1 and X_2 cannot be independent.

2.6.6 For clarity, I'll use shorthand notation. Define

$$X = X_1 - \mu_1, \quad Y = X_2 - \mu_2, \quad Z = X_3 - \mu_3,$$

and similarly substitute X, Y, Z for $1, 2, 3$ in all subscripts ($\mu_X = \mu_1$, etc).

$$\begin{aligned} \sigma_{XY} &= E[XY] = E[E[XY|Y,Z]] \\ &= E[YE[X|Y,Z]] \\ &= E[Y(b_Y Y + b_Z Z)] \\ &= b_Y \sigma_Y^2 + b_Z \sigma_{YZ} \end{aligned}$$

The derivation of σ_{XZ} is symmetric, so

$$\begin{aligned} \begin{pmatrix} \sigma_{XY} \\ \sigma_{XZ} \end{pmatrix} &= \begin{pmatrix} \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{YZ} & \sigma_Z^2 \end{pmatrix} \begin{pmatrix} b_Y \\ b_Z \end{pmatrix} \\ \Rightarrow \begin{pmatrix} b_Y \\ b_Z \end{pmatrix} &= \frac{1}{\sigma_Y^2 \sigma_Z^2 - \sigma_{YZ}^2} \begin{pmatrix} \sigma_Z^2 & -\sigma_{YZ} \\ -\sigma_{YZ} & \sigma_Y^2 \end{pmatrix} \begin{pmatrix} \sigma_{XY} \\ \sigma_{XZ} \end{pmatrix} \\ \Rightarrow b_Y &= \frac{\sigma_{XY}\sigma_Z^2 - \sigma_{XZ}\sigma_{YZ}}{\sigma_Y^2 \sigma_Z^2 (1 - \rho_{YZ}^2)} \\ &= \frac{\sigma_X(\rho_{XY} - \rho_{XZ}\rho_{YZ})}{\sigma_Y(1 - \rho_{YZ}^2)} \end{aligned}$$

And symmetrically for b_Z, so in the notation of the problem,

$$\begin{aligned} b_2 &= \frac{\sigma_1(\rho_{12} - \rho_{13}\rho_{23})}{\sigma_2(1 - \rho_{23}^2)} \\ b_3 &= \frac{\sigma_1(\rho_{13} - \rho_{12}\rho_{23})}{\sigma_3(1 - \rho_{23}^2)} \end{aligned}$$

1. I will assume X, Y are continuous. The proofs are similar for discrete variables, with summations instead of integrals.

$$E\left[E[Y|X]\right] = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y f_{y|x}(y|x) dy \right] f_x(x) dx$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{y|x}(y|x) f_x(x) dy dx$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y) dy dx = E[Y]$$

$$\begin{aligned}
Var[Y] - Var[E[Y|X]] &= \left(E[Y^2] - E[Y]^2\right) - \left(E[E[Y|X]^2] - E[E[Y|X]]^2\right) \\
&= E\Big[E[Y^2|X]\Big] - E\Big[E[Y|X]\Big]^2 - E\Big[E[Y|X]^2\Big] + E\Big[E[Y|X]\Big]^2 \\
&= E\Big[E[Y^2|X] - E[Y|X]^2\Big] \\
&= E\Big[Var[Y|X]\Big] \geq 0 \\
\Rightarrow Var[Y] &\geq Var[E[Y|X]]
\end{aligned}$$

2. (a) The $(i,j)^{th}$ term of the $Var(X)$ matrix is

$$Var(X)_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] =$$
$$Cov(X_i, X_j) = E(X_i X_j) - \mu_i \mu_j$$

Aggregating these terms yields the desired result.

(b)

$$E(AX) = E\left(\sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij} X_j\right) e_i\right)$$
$$= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij} E(X_j)\right) e_i = AE(X) = A\mu$$

For the variance result, first note

$$E[(AX)(AX)'] = E[AXX'A'] = AE[X^2 A']$$
$$= AE[AX^2]' = A[AE[X^2]]' = AE[X^2]A'$$

and

$$E(AX)[E(AX)]' = AE(X)[E(X)]'A' = A[E(X)]^2 A'$$

Combining these results, we get

$$\begin{aligned}Var(AX) &= E[(AX)(AX)'] - E(AX)[E(AX)]' \\ &= A(E[X^2] - E[X]^2)A' = AVar(X)A'\end{aligned}$$

(c) The symmetry of $Var(X)$ follows immediately from the definition, since $E[(X_j - \mu_j)(X_i - \mu_i)] = [(X_i - \mu_i)(X_j - \mu_j)]$. If we let $A = a'$, then (b) tells us

$$a'Var(X)a = AVar(X)A' = Var(AX) = Var(a'X) \geq 0,$$

with the last inequality following because $a'X$ is a scalar random variable.