KEY

Stat 447 Exam 2

April 4, 2002

Prof. Vardeman

1. Below is a table giving a joint probability function for discrete random variables X and Y.

$y \backslash x$	3	4	5	6
4	.1	.1	0	0
3	.05	.2	.05	0
2	0	.05	.2	.05
1	0	0	.1	.1
	.15	.35	.35	.15

7pts a) Evaluate $P[X=Y]$.

$P[X=Y] = p(3,3) + p(4,4) = .05 + .1 = .15$

7pts b) Find the marginal probability function for X, $p(x)$.

x	$p(x)$
3	.15
4	.35
5	.35
6	.15

7pts c) Find the conditional mean of Y given $X=4$, $E[Y \mid X=4]$.

y	$p(y\|4)$	$y\,p(y\|4)$
1	$0/.35 = 0$	0
2	$.05/.35 = \frac{1}{7}$	$2/7$
3	$.2/.35 = \frac{4}{7}$	$12/7$
4	$.1/.35 = \frac{2}{7}$	$8/7$
		$22/7$

$E[Y \mid X=4] = \frac{22}{7}$

6pts d) Are the random variables X and Y independent? Argue very carefully.

X and Y are dependent. For example,

$p(5,4) = 0$ while $p_X(5) = .35$, $p_Y(4) = .2$

and so $p(5,4) \neq p_X(5)\,p_Y(4)$

2. Suppose that U and Z are independent, $U \sim \text{Uniform}(0,1)$ and $Z \sim \text{Normal}(0,1)$.

7pts a) Find the mean and variance of $X = U + Z$.

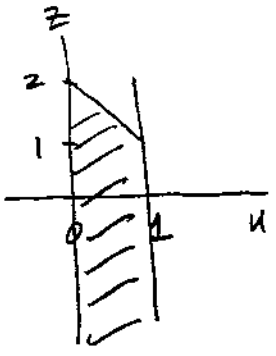
$$E(U+Z) = EU + EZ = \frac{1}{2} + 0 = \frac{1}{2}$$

$$V(U+Z) = (1)^2 V(U) + (1)^2 V(Z) + 2(1)(1)\,\text{cov}(U,Z)$$

$$= \frac{1}{12} + 1 + 0 = \frac{13}{12}$$

$EX =$ $\frac{1}{2}$ $\qquad$ $V(X) =$ $\frac{13}{12}$

9pts b) Completely set up (but do not try to evaluate) a double integral giving $P[U + Z < 2]$. (Use the abbreviation $\phi(z)$ for the standard normal pdf rather than writing out its formula.)



$$\iint 1 \cdot \phi(z)\,du\,dz = \int_0^1 \int_{-\infty}^{2-u} \phi(z)\,dz\,du$$

7pts c) Find the correlation between $X = U + Z$ and $Y = UZ$.

$EY = EU\,EZ = 0$ $\quad$ So $V(Y) = EY^2 = E(UZ)^2 = EU^2Z^2$

independence $\longrightarrow$ $= EU^2\,EZ^2 = \left[\frac{1}{12} + \left(\frac{1}{2}\right)^2\right]\left[1 + 0^2\right] = \frac{1}{3}$

$$\text{cov}(X,Y) = EXY - EX\,EY = EXY$$

$$= E(U+Z)(UZ) = EU^2Z + EUZ^2$$

independence $\Rightarrow$ $EU^2\,EZ + EU\,EZ^2 = 0 + \frac{1}{2}(1) = \frac{1}{2}$

So correlation between X and Y is $\rho = \dfrac{\frac{1}{2}}{\sqrt{\left(\frac{13}{12}\right)\left(\frac{1}{3}\right)}}$

3. Suppose that in describing waiting times to "hits" on a certain web page, it is sensible (in appropriate units) to model

$X_1 =$ time till first hit

$X_2 =$ time between first and second hits and

$X_3 =$ time between second and third hits

as independent $\text{Exp}(1)$ random variables.

6pts a) Find the probability that the time till the first hit exceeds its mean, $P[X_1 > 1]$.

$$P[X_1 > 1] = 1 - P[X_1 \le 1] = 1 - \left(1 - e^{-(1)(1)}\right) = e^{-1} = .3679$$

7pts b) It turns out that $W = X_1/(X_1 + X_2 + X_3)$ has the $\text{Beta}(1,2)$ distribution. (Don't try to show this.) Use this fact to find $P[X_1 > X_2 + X_3]$.

$$P[X_1 > X_2 + X_3] = P[2X_1 > X_1 + X_2 + X_3] = P[W > \tfrac{1}{2}]$$

$$= \int_{\frac{1}{2}}^{1} w^{1-1}(1-w)^{2-1} \frac{1}{B(1,2)}\, dw = \frac{1}{B(1,2)} - \frac{(1-w)^2}{2}\Big|_{\frac{1}{2}}^{1}$$

$$= \frac{1}{8}\frac{1}{B(1,2)} = \frac{1}{4}$$

$$\left(\text{since } B(1,2) = \int_0^1 (1-w)^1\, dw = \tfrac{1}{2}\right)$$

7pts c) Write out (but don't evaluate) an integral giving the expected value of the random variable $Y = \max(X_1, X_2, X_3)$.

Y has pdf $f_{(n)}(y) = n\binom{n}{n} f(y)\left[F(y)\right]^{n-1}$

$$= \begin{cases} 3e^{-y}\left[1 - e^{-y}\right]^2 & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$EY = \int_0^{\infty} y\,(3)\,e^{-y}\left[1 - e^{-y}\right]^2 dy$$

4. Suppose that we model "true" IQs, X, as Normal with $\mu_X = 100$ and $\sigma_X = 10$, and that when an IQ is measured, what is actually observed is $Y = X + \varepsilon$ where the measurement error ε is Normal with $\mu_\varepsilon = 0$ and $\sigma_\varepsilon = 3$. A subject is selected and is tested twice, producing measured IQs $Y_1 = X + \varepsilon_1$ and $Y_2 = X + \varepsilon_2$. Suppose that X, ε_1 and ε_2 are independent.

10pts a) What is the joint distribution of Y_1 and Y_2? $\begin{pmatrix} X \\ \epsilon_1 \\ \epsilon_2 \end{pmatrix} \sim MVN_3\left(\begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 100 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}\right)$

$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ \epsilon_1 \\ \epsilon_2 \end{pmatrix}$ so has a bivariate normal dsn with mean vector $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \end{pmatrix}$ and covariance matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 100 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 100 & 9 & 0 \\ 100 & 0 & 9 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 109 & 100 \\ 100 & 109 \end{pmatrix}$$

(These can easily be worked out <u>without</u> using matrices)

6pts b) What is the correlation between Y_1 and Y_2?

$$\rho = \frac{100}{\sqrt{(109)(109)}} = .9174$$

6pts c) What is the distribution of $\frac{1}{2}(Y_1 + Y_2)$?

$\bar{Y} = \left(\frac{1}{2}, \frac{1}{2}\right)\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ so $\bar{Y}$ is normal with mean $\left(\frac{1}{2}, \frac{1}{2}\right)\begin{pmatrix} 100 \\ 100 \end{pmatrix} = 100$

and variance $\left(\frac{1}{2}, \frac{1}{2}\right)\begin{pmatrix} 109 & 100 \\ 100 & 109 \end{pmatrix}\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = (104.5, 104.5)\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = 104.5$

(This too doesn't really require matrices! $\bar{Y} = X + \bar{\epsilon}$)

8pts 5. Consider the cdf F below. If X has cdf F, find a pdf for $Y = \exp(X)$.

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2}x^2 & \text{if } 0 \le x \le 1 \\ \frac{1}{2}x & \text{if } 1 < x \le 2 \\ 1 & \text{if } 2 < x \end{cases}$$

Note that $f(x) = \begin{cases} x & 0 < x < 1 \\ \frac{1}{2} & 1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$

note also that $g(x) = \exp(x)$ has inverse $g^{-1}(y) = \ln(y)$

So $Y = \exp(X)$ has pdf $f_Y(y) = f(\ln(y))\left|\frac{1}{y}\right|$. This is

$$f_Y(y) = \begin{cases} \ln(y) \cdot \frac{1}{|y|} & 0 < \ln y < 1 \\ \frac{1}{2} \cdot \frac{1}{|y|} & 1 < \ln y < 2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{\ln(y)}{y} & 1 < y < e \\ \frac{1}{2y} & e < y < e^2 \\ 0 & \text{otherwise} \end{cases}$$