MTH135/STA104: Probability

Homework # 10 Due: Tuesday, Dec 6, 2005

Prof. Robert Wolpert

1. Three subjects in a medical trial are given drug A. After one week, those that do not respond favorably are switched to drug B. At the end of the second week, the medical trial is ended and we report

$$X = \text{the number of subjects who respond favorably to drug } A$$
$$Y = \text{the TOTAL number of subjects who respond favorably (to either drug)}$$

If subjects' responses to the drugs are all independent, and if the probability of a favorable response is 1/3 for each subject and each drug, give (either numerically to four correct digits or as a **simplified** formula):

a) $p(x, y)$, the joint PMF for X and Y;

$$\begin{aligned} p(x,y) &= \binom{3}{x} \frac{1}{3}^x \frac{2}{3}^{3-x} \binom{3-x}{y-x} \frac{1}{3}^{y-x} \frac{2}{3}^{3-y} \\ &= \frac{2^{7-x-y}\ \ 3^{x-5}}{x!\,(y-x)!\,(3-y)!}, \qquad 0 \le x \le y \le 3 \end{aligned}$$

or, numerically in tabular form,

$x \backslash y$	0	1	2	3
0	0.0878	0.1317	0.0658	0.0110
1	0	0.1975	0.1975	0.0494
2	0	0	0.1481	0.0741
3	0	0	0	0.0370

b) $p_1(x)$, the marginal PMF for X;

$$p_1(x) \;=\; \binom{3}{x}\frac{1}{3}^x\,\frac{2}{3}^{3-x} \;=\; \frac{2^{4-x}}{9\,x!\,(3-x)!}, \qquad 0 \le x \le 3$$

or, numerically in tabular form,

x	0	1	2	3
$p(x)$	0.2963	0.4444	0.2222	0.0370

c) $p_2(y)$, the marginal PMF for Y;

Each subject has two opportunities for success, with total success probability $p = 1 - \frac{2}{3}^2 = \frac{5}{9}$; thus

$$p_2(y) \;=\; \binom{3}{y}\frac{5}{9}^y\,\frac{4}{9}^{3-y} \;=\; \frac{128\,(5/4)^y}{243\,y!\,(3-y)!}, \qquad 0 \le y \le 3$$

or, numerically in tabular form,

y	0	1	2	3
$p(y)$	0.0878	0.3292	0.4115	0.1715

d) $p(y|x=2)$, the conditional PMF for Y given $X = 2$;

e) $p(x|y=2)$, the conditional PMF for X given $Y = 2$.

Equivilently, roll three fair six-sided dice and count the number X of divisible-by-three (i.e., 3 or 6) rolls; roll any *in*divisible-by-three dice again, leading to a total altogether of Y dice-divisible-by-three.

2. A drawer has ten pairs of socks— five pairs of white athletic socks, three pairs of black socks, and two pairs of pretty purple paisley-pattern socks. Unfortunately the twenty socks are all tossed in at random, instead of being nicely paired like Mom would have done. We reach in blindly and pull out three socks at random, without replacement. Let W be the number of white socks drawn, B the number of black ones, and P the number of pretty ones. One can show (you don't have to) that

$$\mathsf{P}[W = w, B = b, P = p] = \frac{\binom{10}{w}\binom{6}{b}\binom{4}{p}}{\binom{20}{3}}$$

for any three integers $w \geq 0$, $b \geq 0$, $p \geq 0$ with $w+b+p = 3$, while

$$\mathsf{P}[W = w] = \frac{\binom{10}{w}\binom{10}{3-w}}{\binom{20}{3}} \quad \mathsf{P}[B = b] = \frac{\binom{6}{b}\binom{14}{3-b}}{\binom{20}{3}} \quad \mathsf{P}[P = p] = \frac{\binom{4}{p}\binom{16}{3-p}}{\binom{20}{3}}$$

a) What is the (marginal) expectation $\mathsf{E}[B]$? Note this is just a number.

Each of the three socks pulled has probability $6/20$ of being black, giving an expectation of $\mathsf{E}[B] = 9/10$. B has a hypergeometric $\mathsf{HG}(3, 6, 14)$ distribution.

b) What is the (conditional) expectation $\mathsf{E}[B|W]$? Note this is a function of w; for which values of w must it be defined?

The possible numbers of white socks pulled are $w \in \{0, 1, 2, 3\}$. The conditional distribution of B given W is given by

$$\mathsf{P}[B = b \mid W = w] = \frac{\binom{10}{w}\binom{6}{b}\binom{4}{3-w-b}/\binom{20}{3}}{\binom{10}{w}\binom{10}{3-w}/\binom{20}{3}} = \frac{\binom{6}{b}\binom{4}{3-w-b}}{\binom{10}{3-w}},$$

so evidently B has a conditional hypergeometric $\mathsf{HG}(3-w, 6, 4)$ distribution with mean $\mathsf{E}[B \mid W] = (3-w) \times \frac{6}{10} = (18-6w)/10$.

Alternatively— 60% of the non-white socks are black, so the expected number of black socks among the $(3-w)$ non-white ones is

$$\mathsf{E}[B \mid W] = \frac{6\,(3-w)}{10} = \frac{18 - 6\,w}{10}, \qquad w = 0, \cdots, 3.$$

c) Give the conditional probability $\mathsf{P}[B = 1 \mid W = 1]$. Extra credit for identifying (by name, with correct parameters) the conditional distribution $\mathsf{P}[B = b \mid W = 1]$ of B, given $W = 1$.

From above the conditional distribution of B given $W = 1$ is $\mathsf{HG}(2, 6, 4)$ and hence

$$\mathsf{P}[B = 1 \mid W = 1] = \frac{\binom{6}{1}\binom{4}{3-1-1}}{\binom{10}{3-1}} = \frac{6 \times 4}{45} = \frac{8}{15} \approx 0.5333;$$

alternatively we can compute the needed probabilities directly:

$$\begin{aligned}
\mathsf{P}[W = 1, B = 1] &= 3! \times \frac{10}{20} \times \frac{6}{19} \times \frac{4}{18} = \frac{4}{19} \approx 0.2105 \\
\mathsf{P}[W = 1] &= 3 \times \frac{10}{20} \times \frac{10}{19} \times \frac{9}{18} = \frac{15}{38} \approx 0.3947 \\
\mathsf{P}[B = 1 \mid W = 1] &= \mathsf{P}[W = 1, B = 1]/\mathsf{P}[B = 1] \\
&= \frac{4/19}{15/38} = \frac{8}{15} \approx 0.5333
\end{aligned}$$

3. First the random variable $X \sim \mathsf{Un}(0, 10)$ is chosen uniformly on the interval from zero to ten; then $Y \sim \mathsf{Un}(0, X)$ is chosen uniformly (think of breaking a 10-cm piece of chalk at random, then breaking the resulting piece at random. I do this most days in lecture...).

a) Find the joint density function $f(x, y)$ for X and Y, correctly for every $(x, y) \in \mathbb{R}^2$;

$$f(x, y) = \frac{1}{10\,x}, \qquad 0 < y < x < 10 \quad \text{(otherwise 0)}$$

b) Find both marginal density functions $f_1(x)$ and $f_2(y)$;

$$
\begin{aligned}
f_1(x) &= \int_{-\infty}^{\infty} f(x,y)\,dy \\
&= \int_0^x \frac{1}{10\,x}\,dy \;=\; \frac{1}{10}, \qquad 0 < x < 10 \\
f_2(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \\
&= \int_y^{10} \frac{1}{10\,x}\,dx \;=\; \frac{\ln(10)-\ln(y)}{10} = \frac{\ln(10/y)}{10}, \qquad 0 < y < 10
\end{aligned}
$$

c) Find both conditional density functions $f(x \mid y)$ and $f(y \mid x)$;

$$
\begin{aligned}
f(x|y) &= \frac{1/10x}{\ln(10/y)/10} = \frac{1}{x\,\ln(10/y)}, \qquad y < x < 10 \\
f(y|x) &= \frac{1/10x}{1/10} = \frac{1}{x}, \qquad 0 < y < x
\end{aligned}
$$

d) Find both conditional expectations $\mathsf{E}[X|Y]$ and $\mathsf{E}[Y|X]$.

$$
\begin{aligned}
E[X|y] &= \int_y^{10} \frac{x}{x\,\ln(10/y)}\,dx = \frac{10-y}{\ln 10 - \ln y} \\
E[Y|x] &= \int_0^x \frac{y}{x}\,dy = \frac{x^2}{2\,x} = \frac{x}{2}
\end{aligned}
$$

4. A fair six-sided die is rolled, showing some number X; we then let Y be the sum of X independent normally-distributed random variables, so

$$X \sim \mathsf{Un}\{1, 2, ..., 6\} \qquad Y = \sum_{i=1}^{X} Z_i, \quad Z_i \overset{\text{iid}}{\sim} \mathsf{No}(0, 1).$$

a) What is the conditional distribution of Y, given X?

$$Y \sim \mathsf{No}(\mu = 0,\ f\sigma^2 = X)$$

b) What are the mean and variance of Y? (Hint: Start by finding $\mathsf{E}[Y]$ and $\mathsf{E}[Y^2]$ by computing $\mathsf{E}[Y|X]$ and $\mathsf{E}[Y^2|X]$ first).

$$\begin{aligned} \mathsf{E}[Y|X] &= 0 \\ \mathsf{E}[Y] &= E[0] \ = \ 0 \\ \mathsf{E}\left[Y^2|X\right] &= X \\ \mathsf{E}\left[Y^2\right] &= \mathsf{E}[X] = \frac{7}{2} \\ \mathsf{Var}[Y] &= \mathsf{E}[Y^2] - (\mathsf{E}[Y])^2 = \frac{7}{2} \end{aligned}$$

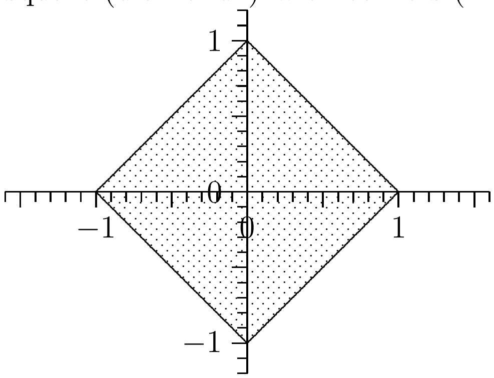
c) Does Y have a normal distribution? Explain.

No. The probability density function for Y is

$$f(y) \ = \ \sum_{x=1}^{6} \frac{1}{6} \frac{1}{\sqrt{2\pi x}} e^{-y^2/2x} \ = \ \frac{1}{6\sqrt{2\pi}} \sum_{x=1}^{6} e^{-y^2/2x} / \sqrt{x},$$

not the normal pdf.

5. The variables X and Y are drawn from the uniform distribution on the square (diamond?) with corners $(\pm 1, 0)$ and $(0, \pm 1)$.



a) The joint density function $f(x, y)$ is some constant $c > 0$ on the set where $|x| + |y| \leq 1$, and zero elsewhere. What is the numerical value of c?

One over the area of the square, or $c = 1/2$.

b) Verify that the marginal densities $f_1(x)$ and $f_2(y)$ are both

$$f(s) = 1 - |s|, \; s \in (-1, 1); \quad f(s) = 0, \; s \notin (-1, 1).$$

$$f_1(x) \;=\; \int_{-\infty}^{\infty} f(x, y)\, dy \;=\; \int_{|x|-1}^{1-|x|} \frac{1}{2}\, dy \;=\; 1 - |x|, \qquad -1 < x < 1.$$

The case for $f_2(y)$ is identical.

c) Find the conditional density function $f(x|y)$ for X, given Y.

$$f(x|y) = \frac{1/2}{1 - |y|} = \frac{1}{2 - 2|y|}, \qquad |y| - 1 < x < 1 - |y|$$

d) Find the conditional expectation $\mathsf{E}[X \mid Y]$.

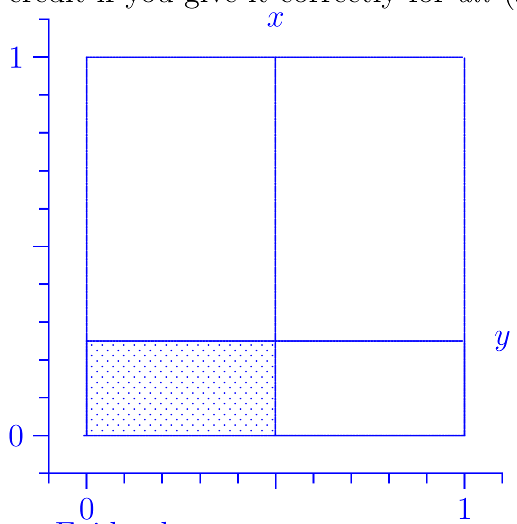
$$\mathsf{E}[X \mid Y] \;=\; \int_{-\infty}^{\infty} x\, f(x|y)\, dx \;=\; \int_{|y|-1}^{1-|y|} \frac{x}{2}\, dx \;=\; 0.$$

e) Are X and Y independent? Why?

No. For example, $f(x, y) = 0$ at $x = y = \frac{3}{4}$, but $f_1(0.75) = f_2(0.75) = 0.25 > 0$.

6. The point (X, Y) is drawn uniformly from the unit square $[0,1] \times [0,1]$ in the plane. We are interested in the conditional distributions of X, the horizontal coordinate, given several different possible conditions.

a) Find the joint distribution function for X and Y at each (x, y) in the set $0 \le x \le 1$, $0 \le y \le 1$. Draw a picture— no integration is needed! Extra credit if you give it correctly for *all* $(x, y) \in \mathbb{R}^2$.



Evidently

$$F(x,y) = \begin{cases} 0 & x \le 0 \text{ or } y \le 0 \\ x\,y & 0 < x \le 1,\ 0 < y \le 1 \\ x & 0 < x \le 1,\ 1 < y < \infty \\ y & 1 < x < \infty,\ 0 < y \le 1 \\ 1 & 1 < x < \infty,\ 1 < y < \infty \end{cases}$$

The joint pdf (needed below) is $f(x, y) = 1$ on $0 < x \le 1$, $0 < y \le 1$.

b) Find the marginal pdf for Y:

Taking the derivitive above w.r.t y as $x \to \infty$,

$$f_2(y) = \begin{cases} 1 & 0 < y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

c) Find the conditional density function $f(x \mid y)$ of X, given Y, for $0 < y \le 1$.

Taking the ratio $f(x, y)/f_2(y)$,

$$f(x \mid y) = \begin{cases} 1 & 0 < x \le 1, \ 0 < y \le 1 \\ 0 & \text{other } x \text{ and } y \end{cases}$$

(more precisely it's undefined for $y \notin (0, 1]$).

d) The marginal mean of X is $\mathsf{E}[X] = 1/2$. What is the conditional mean of X, given $Y = 0$? Find $\mathsf{E}[X|Y]$.

From c) above, the mean is

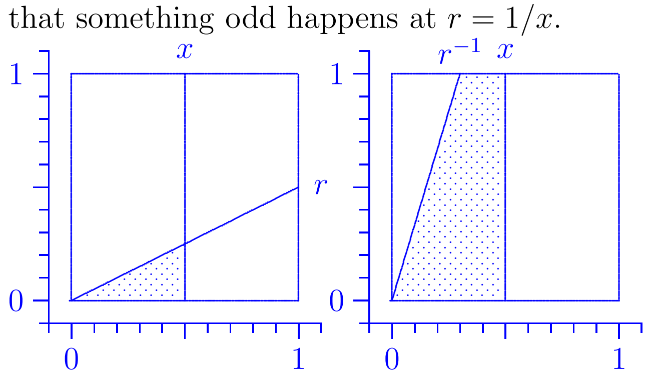
$$\mathsf{E}[X \mid Y] = \frac{1}{2}$$

for any $0 < y \le 1$.

7. The point (X, Y) is (again) drawn uniformly from the unit square $[0, 1] \times [0, 1]$ in the plane. We are still interested in the conditional distributions of X, the horizontal coordinate, given several different possible conditions.

This time, set $R \equiv Y/X$, and:

a) Find the joint distribution function for R and X at each (r, x) in the set $0 \le r < \infty$, $0 \le x \le 1$. Draw a picture— no integration is needed! Note that something odd happens at $r = 1/x$.



Evidently

$$F(r,x) = \begin{cases} 0 & x < 0 \text{ or } r < 0 \\ r\,x^2/2 & 0 < x < 1,\ 0 < r < 1/x \\ x - 1/2r & 0 < x < 1,\ 1/x < r < \infty \\ r/2 & x > 1,\ 0 < r < 1 \\ 1 - 1/2r & x > 1,\ 1 < r < \infty \end{cases}$$

The joint pdf (needed below) is $f(r,x) = x$ on $0 < x < 1$, $0 < r < 1/x$.

b) Find the marginal pdf for R:

Taking the derivitive above w.r.t r as $x \to \infty$,

$$f_1(r) = \begin{cases} 0 & r < 0 \\ 1/2 & 0 < r < 1 \\ 1/2r^2 & 1 < r < \infty \end{cases}$$

c) Find the conditional density function $f(x \mid r)$ of X, given R, for $0 < r \le 1$. Extra credit for $0 < r < \infty$.

Taking the ratio $f(r,x)/f_1(r)$,

$$f(x \mid r) = \begin{cases} 2x & 0 < x < 1,\ 0 < r \le 1 \\ 2\,x\,r^2 & 0 < x < 1/r,\ 1 < r < \infty \\ 0 & \text{other } x \text{ and } r \end{cases}$$

d) The marginal mean of X is $\mathsf{E}[X] = 1/2$. What is the conditional mean of X, given $R = 0$? Given $R = 1$? Find $\mathsf{E}[X|R]$.

From c) above, the mean is

$$\mathsf{E}[X \mid R] = \begin{cases} \frac{2}{3} & 0 < r \le 1 \\ \frac{2}{3r} & 1 < r < \infty, \end{cases}$$

or $2/3$ at $r = 0$ and also at $r = 1$.

8. If all went well, you got different answers for the last two problems. This might seem mysterious to you (I hope so)— since in each case, somehow we're asking about how big X would be, on average, if (X, Y) lies in the same set

$$\Delta = \left\{(x,y) : 0 < x < 1,\ R \equiv \frac{y}{x} = 0\right\} = \{(x,y) : 0 < x < 1,\ y = 0\}$$

Unfortunately Δ has zero area, so $\mathsf{P}[(X,Y) \in \Delta] = 0$ and the notation "$\mathsf{E}[X \mid (X,Y) \in \Delta]$" isn't well-defined (that's the point of these problems). Here's something to help explain this "paradox."

For any set A with positive probability $\mathsf{P}[(X,Y) \in A] > 0$ we calculate

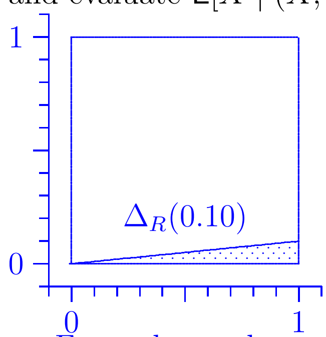
$$\mathsf{P}[X \le x \mid (X,Y) \in A] = \frac{\mathsf{P}[X \le x \ \cap\ (X,Y) \in A]}{\mathsf{P}[(X,Y) \in A]}$$

in the usual way and, from this, we can find the conditional distribution of X and the conditional expectation of X, given $(X,Y) \in A$. Sketch the regions described below and find the conditional expectations requested:

a) For $\epsilon = 0.10$ and $R \equiv y/x$, sketch and shade the set

$$\Delta_R(\epsilon) \equiv \{(x,y) : 0 < x < 1,\ 0 < R < \epsilon\}$$

and evaluate $\mathsf{E}[X \mid (X,Y) \in \Delta_R(\epsilon)]$.



From above, the conditional CDF and pdf are

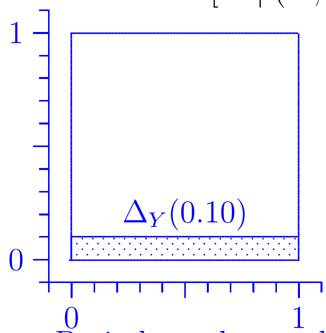
$$\begin{aligned} F\big(x \mid (X,Y) \in \Delta_R(\epsilon)\big) &= \frac{\epsilon\, x^2/2}{\epsilon/2} = x^2 \\ f\Big(x \mid (X,Y) \in \Delta_R(\epsilon)\Big) &= 2x \end{aligned}$$

on the set $0 < x < 1$ and hence X has a $\mathsf{Be}(2,1)$ conditional distribution with mean $\mathsf{E}[X \mid (X,Y) \in \Delta_R(\epsilon)] = \frac{2}{3}$.

b) For $\epsilon = 0.10$, sketch and shade the set

$$\Delta_Y(\epsilon) \equiv \{(x, y) : 0 < x < 1,\ 0 < y < \epsilon\}$$

and evaluate $\mathsf{E}[X \mid (X, Y) \in \Delta_Y(\epsilon)]$.



By independence, the conditional CDF and pdf are

$$\begin{aligned} F(x \mid (X, Y) \in \Delta_Y(\epsilon)) &= \frac{\epsilon\, x}{\epsilon} = x \\ f(x \mid (X, Y) \in \Delta_Y(\epsilon)) &= 1 \end{aligned}$$

on the set $0 < x < 1$ and hence X has a uniform conditional distribution with mean $\mathsf{E}[X \mid (X, Y) \in \Delta_Y(\epsilon)] = \frac{1}{2}$.

c) Explain from your pictures above which of the following you should expect. Were you right?

$$\begin{aligned} \text{a) } \mathsf{E}[X \mid R = 0] &< \mathsf{E}[X \mid Y = 0] \\ \text{b) } \mathsf{E}[X \mid R = 0] &= \mathsf{E}[X \mid Y = 0] \\ \text{c) } \mathsf{E}[X \mid R = 0] &> \mathsf{E}[X \mid Y = 0] \end{aligned}$$

c) $\mathsf{E}[X \mid R = 0] > \mathsf{E}[X \mid Y = 0]$, since X tends to be bigger in the wedge $\Delta_R(\epsilon)$ than in the slab $\Delta_Y(\epsilon)$. It *is* right, since $\frac{2}{3} > \frac{1}{2}$.