Homework Set 9 *DRAFT* **EECS 401** **Due: April 3, 2000,**

1. *Joe buys his gasoline at the local cut-rate gas station. The station sells two grades of gas, A and B, but the station owner won't tell which you get. Joe is concerned because his car mileage varies uniformly between 20 and 32 miles per gallon with brand A and between 16 and 26 miles per gallon with brand B. Suppose Joe knows the probabilities of the gas station having brands A and B are .3 and .7, respectively. Let X be a random variable representing Joe's gas mileage. (This is same descriptions as Prob. 5 of previous homework.)*

   Find the expected value of X WITHOUT integrating x times the density of X.

   We use a version of the law of total expectation, namely,

   $$E[X] = E[X|A]\,P(A) + E[X|B]\,P(B)$$

   From the problem statement, we deduce Pr(A) =0.3, Pr(B)=0.7; and

   $$f_X(x|A) = \begin{cases} \frac{1}{12}, & 20 \le x \le 32 \\ 0, & \text{else} \end{cases} \qquad f_X(x|B) = \begin{cases} \frac{1}{10}, & 16 \le x \le 26. \\ 0, & \text{else} \end{cases}$$

   Hence, $E[X|A] = \int_{20}^{32} \frac{1}{12}\,dx = 26$ and $E[X|B] = \int_{-16}^{26} \frac{1}{10}\,dx = 21$, and

   $$\mathbf{E[X]} = 26 \times .3 + 21 \times .7 = \mathbf{22.5}$$

2. *A gambler brings X dollars to a casino where X is a random variable with density*

   $$f_X(x) = \begin{cases} \frac{x}{80{,}000}, & 0 \le x \le 400 \\ 0, & \text{otherwise} \end{cases}$$

   After a night of gambling the gambler leaves the casino with Y dollars, where Y is uniformly distributed between 0 and X . (This is same descriptions as Prob. 6 of previous homework.)

   (a) *Find the expected value of Y WITHOUT integrating x times the density of Y.*

   We use a version of the law of total expectation, namely, $E[Y] = \int_{-\infty}^{\infty} E[Y|X=x]\, f_X(x)\,dx$

   From the problem statement, we deduce $f_{Y|X}(y|x) = \begin{cases} \frac{1}{x}, & 0 \le y \le x \\ 0, & \text{else} \end{cases}$

   Therefore, $E[Y|X=x] = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\,dy = \int_{0}^{x} y\,\frac{1}{x}\,dy = \frac{1}{x}\frac{y^2}{2}\Big|_0^x = \frac{x}{2}$

   $$\Rightarrow \qquad \mathbf{E[Y]} = \int_0^{400} \frac{x}{2}\,\frac{x}{80{,}000}\,dx = \frac{1}{160{,}000}\frac{x^3}{3}\Big|_0^{400} = \frac{400}{3} = \mathbf{133.3}$$

   (a) *Given the gambler leaves the casino with $200 dollars, find the expected number of dollars that he brought.*

   We need to find $E[X|Y=200] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|200)\,dx$

   To find $f_{X|Y}(x|y)$ we use the following version of Bayes rule: $f_{X|Y}(x|200) = \frac{f_{Y|X}(200|x)\, f_X(x)}{f_Y(200)}$

where $f_Y(200) = \int_{-\infty}^{\infty} f_X(x)\, f_{Y|X}(200|x)\, dx = \int_{200}^{400} \frac{x}{80{,}000} \frac{1}{x} dx = \frac{200}{80{,}000} = \frac{1}{400}$

(note that $f_{Y|X}(y|x) = \frac{x}{80{,}000}$ only if $0 \le y \le x$)

So $f_{X|Y}(x|200) = \dfrac{\frac{1}{x}\frac{x}{80{,}000}}{\frac{1}{400}} = \frac{1}{200}$ if $200 \le x \le 400$

And $\mathbf{E[X|Y=200]} = \int_{200}^{400} x \frac{1}{200} dx = \frac{1}{200} \frac{x^2}{2}\Big|_{200}^{400} = 400 - 100 = \mathbf{300}$

3. 4.39, p. 260

$f_{X,Y,Z}(x,y,z) = k(x+y+z), \quad 0 \le x \le 1,\ 0 \le y \le 1,\ 0 \le z \le 1$

(a) We choose k so that $f_{X,Y,Z}(x,y,z)$ integrates to one.

$$1 = \int_0^1 \int_0^1 \int_0^1 f_{X,Y,Z}(x,y,z)\, dz\, dy\, dx = \int_0^1 \int_0^1 \int_0^1 k\,(x+y+z)\, dz\, dy\, dx$$

$$= \int_0^1 \int_0^1 k\,(x+y+\tfrac{1}{2})\, dy\, dx = \int_0^1 k\,(x+\tfrac{1}{2}+\tfrac{1}{2})\, dx = k\left(\tfrac{1}{2}+\tfrac{1}{2}+\tfrac{1}{2}\right) = \tfrac{3}{2}k \Rightarrow \mathbf{k = \tfrac{2}{3}}$$

(b) $f_{X,Y}(x,y) = \int_0^1 f_{X,Y,Z}(x,y,z)\, dz = \frac{2}{3}(x+y+\frac{1}{2}), \ 0 \le x \le 1,\ 0 \le y \le 1$

$$\mathbf{f_{Z|X,Y}(z|x,y)} = \frac{f_{X,Y,Z}(x,y,z)}{f_{X,Y}(x,y)} = \frac{\frac{2}{3}(x+y+z)}{\frac{2}{3}(x+y+\frac{1}{2})} = \mathbf{\frac{(x+y+z)}{(x+y+\frac{1}{2})}}$$

4. 4.41, p. 261 *(don't use the chain rule; you are, in effect, rederiving it for 3 random variables.)*

We need to show $f_{X,Y,Z}(x,y,z) = f_X(x)\, f_{Y|X}(y|x)\, f_{Z|X,Y}(z|x,y)$. Let us start with the right hand side and substitute the formulas for each of the conditional densities there. We have

$$f_X(x)\, f_{Y|X}(y|x)\, f_{Z|X,Y}(z|x,y) = f_X(x) \frac{f_{XY}(x,y)}{f_X(x)} \frac{f_{XYZ}(x,y,z)}{f_{XY}(x,y)} = f_{XYZ}(x,y,z)$$

Alternate derivation: We use the fact that $f_{U,V}(u,v) = f_U(u)\, f_{V|U}(v|u)$

With $U = (X,Y)$, $u = (x,y)$, $V = Z$ and $v = z$, the above fact implies

$$f_{X,Y,Z}(x,y,z) = f_{U,V}(u,v) = f_U(u)\, f_{V|U}(v|u) = f_{X,Y}(x,y)\, f_{Z|X,Y}(z|x,y)$$
$$= f_X(x)\, f_{Y|X}(y|x)\, f_{Z|X,Y}(z|x,y)$$

where the last step follows by the above stated fact, with $U = X$, $u = x$, $V = Y$, $v = y$

5. *Let $W = X + 3Y - Z$, where X, Y and Z be uncorrelated Gaussian random variables with means $m_x = 1$, $m_Y = 2$ and $m_Z = 3$, and variances $\sigma_X^2 = 1$, $\sigma_Y^2 = 2$, $\sigma_Z^2 = 3$.*

Find $P(W \ge 1)$. Hint: W is a linear combination of Gaussian random variables.

W is Gaussian, since it is the linear combination of Gaussian random variables. Therefore, to find the density of W we need only find E[W] and σ_W^2.

$$E[W] = E[X+3Y-Z] = E[X] + 3\,E[Y] - E[Z] = 1 + 3\times 2 - 3 = 4$$

Since X, Y and Z are uncorrelated, so are X, 3Y and -Z, so the variance of W = X + 3Y - Z is the sum of the variances of X, 3Y and -Z:

$$\sigma_W^2 = \text{var}(X) + \text{var}(3Y) + \text{var}(-Z) = \sigma_X^2 + 9\sigma_Y^2 + \sigma_Z^2 = 1 + 9\times 2 + 3 = 22$$

Now,

$$\mathbf{P(W\geq 1)} = \int_1^\infty f_W(w)\,dw = \int_1^\infty \frac{1}{\sqrt{2\pi\sigma_W^2}} \exp\left\{-\frac{(w-E[W])^2}{2\sigma_W^2}\right\} dw$$

$$= \int_{\frac{1-E[W]}{\sigma_W}} \frac{1}{\sqrt{2\pi\sigma_W^2}} \exp\left\{-\frac{u^2}{2}\right\} du\,\sigma_W, \quad \text{letting } u = \frac{w-E[W]}{\sigma_W}$$

$$= Q\left(\frac{1-E[W]}{\sigma_W}\right) = Q\left(\frac{1-4}{\sqrt{22}}\right) = Q(-.64) = 1 - Q(.64) \cong 1 - .26 \cong \mathbf{.74}$$

6. 4.51, p. 262

$f_{X,Y}(x,y) = 2e^{-(x+y)}$, $0 \leq y \leq x$, and Z = X + Y. Note that X and Y are not independent. As shown in class and in Example 4.31, p. 221,

$$\mathbf{f_Z(z)} = \int_{-\infty}^{\infty} f_{XY}(x,z-x)\,dx, \qquad \text{note: } f_{XY}(x,z-x) = 2e^{-(x+z-x)}, \; 0 \leq z-x \leq x$$

$$= 2\,e^{-z}, \quad z/2 \leq x \leq z \;=\; \int_{z/2}^{z} 2\,e^{-z}\,dx = 2\,e^{-z}\,\frac{z}{2} = \mathbf{z\,e^{-z}}, \quad \mathbf{z \geq 0}$$

7. 4.61, p. 263

Since X and Y are independent X^2 and Y are independent (functions of independent random variables are independent), so

$$E[X^2 Y] = E[X^2]\;E[Y]$$

Since X is Gaussian with E[X] = 0 and $\sigma_X^2 = 1$, $E[X^2] = \sigma_X^2 + (E[X])^2 = 1 + 0 = 1$

Since Y is uniform on [-1,3], E[Y] = 1.

$$\mathbf{E[X^2 Y]} = E[X^2]\;E[Y] = 1 \times 1 = \mathbf{1}.$$

8. 4.67, p. 263

$Y = aX + b$ and $\rho = \dfrac{\text{cov}(X,Y)}{\sigma_X\sigma_Y}$,

$$\text{cov}(X,Y) = E[XY] - E[X]\,E[Y] = E[X(aX+b] - E[X]\,E[aX+b]$$

$$= a\,E[X^2] + b\,E[X] - a\,(E[X])^2 - bE[X] = a\,(E[X^2] - (E[X])^2) = a\,\sigma_X^2$$

$$\sigma_Y^2 = \text{var}(aX+b) = a^2 \sigma_X^2 \quad \Rightarrow \quad \rho = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{a\, \sigma_X^2}{\sigma_X\, a\, \sigma_X} = \mathbf{1}$$

9. *Let X,Y be continuous random variables with joint density*

$$f_{XY}(x,y) = \begin{cases} x+y, & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{else} \end{cases}$$

(a) *Find the best linear estimate for Y given $X=x$.*

The best linear estimate for Y given $X = x$, is

$$\hat{y} = ax + b = \rho_{XY} \frac{\sigma_Y}{\sigma_X} (X - E[X]) + E[Y]$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx = \int_0^1 (x+y)\, dx = \left(\frac{x^2}{2}\Big|_0^1 + y \right) = \frac{1}{2} + y, \; 0 < y < 1$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy = \int_0^1 (x+y)\, dy = \left(\frac{y^2}{2}\Big|_0^1 + x \right) = \frac{1}{2} + x, \; 0 < x < 1$$

$$E[X] = E[Y] = \int_0^1 y \left(\frac{1}{2} + y \right) dy = \frac{1}{4} + \frac{1}{3} = \frac{7}{12}$$

$$E[X^2] = E[Y^2] = \int_0^1 y^2 \left(\frac{1}{2} + y\right) dy = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}$$

$$E[XY] = \int_0^1 \int_0^1 xy\,(x+y)\, dx\, dy = \int_0^1 \frac{y}{3} + \frac{y^2}{2}\, dy = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$\rho_{XY} = \frac{E[XY]-E[X]E[Y]}{\sqrt{E[X^2]-(E[X])^2}\sqrt{E[Y^2]-(E[Y[])^2}} = \frac{1/3 - 49/144}{5/12-49/144} = -\frac{1}{11}$$

So, $\hat{\mathbf{y}} = \mathbf{-\frac{1}{11}\left(x - \frac{7}{12}\right) + \frac{7}{12}}$

(b) *Find the best overall estimate for Y given $X = x$.*

The best overall estimate for Y given $X=x$ is $\hat{y} = E[Y|X=x] = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dx$

Using Bayes rule: $f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{x+y}{1/2+x} = \frac{2(x+y)}{1+2x}$, $0 < x < 1$, $0 < y < 1$

Therefore, $\hat{\mathbf{y}} = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dx = \int_0^1 y\, \frac{2(x+y)}{1+2x}\, dy = \frac{2x}{1+2x} \int_0^1 y\, dy + \frac{2}{1+2x} \int_0^1 y^2\, dy$

$$= \frac{2x}{1+2} \frac{y^2}{2}\Big|_0^1 + \frac{2}{1+2x} \frac{y^3}{3}\Big|_0^1$$

$$= \mathbf{\frac{2}{1+2x} \left(\frac{x}{2} + \frac{1}{3}\right)}$$

The two estimation rule are plotted $\Rightarrow$
(One is linear, one is not.)