SOLUTION FOR HOMEWORK 9, STAT 4352

Welcome to your homework devoted to regression. Here all problems are devoted to a random design regression. They are good exercises to improve/refresh your technical skills.

As usual, try to find mistakes (and get extra points) in my solutions. Typically they are silly arithmetic mistakes (not methodological ones). They allow me to check that you did your HW on your own. Please do not e-mail me about your findings — just mention them on the first page of your solution and count extra points.

Now let us look at your problems.

1. Problem 14.1. The joint pdf is

$$f^{X,Y}(x,y) = xe^{-x(1+y)}I(x>0)I(y>0).$$

Note that X and Y are dependent (via the factor e^{-xy}).

We need to calculate the marginal pdf

$$f^Y(y) = \int_{-\infty}^{\infty} f^{X,Y}(x,y)dx = \int_0^{\infty} xe^{-x(1+y)}dxI(y>0).$$

I can calculate this integral using either integration by parts, or Gamma density approach, or recall that for $Z \sim Expon(\theta)$ with $f^Z(z) = \theta^{-1}e^{-z/\theta}I(z>0)$ we have $E(Z) = \theta$. The latter is the simplest approach, and it yields

$$f^Y(y) = (1+y)^{-2}I(y>0).$$

Then for $y > 0$

$$E(X|Y=y) = (1+y)^2 \int_0^{\infty} x^2 e^{-x(1+y)}dx.$$

Now again either Gamma/Exponential approach or integration by parts can be used. Let us recall the latter:

$$\int_0^{\infty} x^2 e^{-x(1+y)}dx = (-1)(1+y)^{-1}x^2e^{-x(1+y)}|_{x=0}^{\infty} + 2(1+y)^{-1}\int_0^{\infty} xe^{-x(1+y)}dx = 2(1+y)^{-1}(1+y)^{-2}.$$

Combining the results we conclude that for $y > 0$

$$E(X|Y=y) = 2(1+y)^{-1}.$$

Note that this is the best (in terms of the minimal mean squared error) regression curve which predicts X for a given $Y = y$.

2. Problem 14.2. Here

$$f^{X,Y}(x,y) = (2/5)(2x+3y)I(0<x<1)I(0<y<1).$$

(a) For $x \in (0,1)$

$$E(Y|X=x) = \int_{-\infty}^{\infty} yf^{Y|X}(y|x)dy.$$

We need the marginal

$$f^X(x) = \int_{-\infty}^{\infty} f^{Y,X}(y,x)dy = \int_0^1 (2/5)(2x+3y)dyI(x \in (0,1))$$

$$= (2/5)\Big[2xy + (3/2)y^2\Big]_{y=0}^1 I(x \in (0,1)) = (2/5)[2x + 3/2]I(0 < x < 1).$$

Then for $x \in (0,1)$

$$E(Y|X=x) = \int_0^1 y\frac{(2/5)(2x+3y)}{(2/5)(2x+3/2)}dy = (2x+3/2)^{-1}\Big[xy^2 + y^3\Big]_{y=0}^1 = \frac{x+1}{2x+3/2}.$$

(b) For $y \in (0,1)$,

$$E(X|Y=y) = \int_{-\infty}^{\infty} x f^{X|Y}(x|y)dx.$$

Let us calculate the marginal density

$$f^Y(y) = \int_{-\infty}^{\infty} f^{Y,X}(y,x)dx = (2/5)\int_0^1 (2x+3y)dxI(y \in (0,1))$$

$$= (2/5)\Big[x^2 + 3yx\Big]_{x=0}^1 I(y \in (0,1)) = (2/5)[1+3y]I(y \in (0,1)).$$

Then for $y \in (0,1)$

$$E(X|Y=y) = \int_0^1 \frac{x(2/5)(2x+3y)}{(2/5)(1+3y)}dx$$

$$= (1+3y)^{-1}\Big[(2/3)x^3 + (3/2)yx^2\Big]_{x=0}^1 = \frac{(2/3)+(3/2)y}{1+3y}.$$

3. Problem 14.3. Here

$$f^{X,Y}(x,y) = 6xI(0 < x < y < 1).$$

Note that X and Y are dependent via the support (the indicator function cannot be written as a product $g_1(x)g_2(y)$). Here the support is the upper triangle $y > x$ in the unit square $[0,1]^2$. Then

$$E(Y|X=x) = \int_{-\infty}^{\infty} y f^{Y|X}(y|x)dy.$$

For the marginal

$$f^X(x) = \int_{-\infty}^{\infty} f^{Y,X}(y,x)dx = \int_0^1 6xI(0 < x < y < 1)dy$$

$$= 6x\int_x^1 dyI(0 < x > 1) = 6x(1-x)I(0 < x < 1).$$

Please check that it is pdf, that is the function is nonnegative and integrated to 1.

Then for $x \in (0,1)$:

$$E(Y|X=x) = \int_x^1 \frac{6yx}{6x(1-x)}dx = \frac{y^2|_x^1}{2(1-x)}$$

$$= \frac{1-x^2}{2(1-x)} = (1+x)/2.$$

(b) For $y \in (0,1)$

$$E(X|Y=y) = \int_{-\infty}^{\infty} x f^{X|Y}(x|y)dx.$$

Write,

$$f^Y(y) = \int_0^y 6xdxI(0<y<1) = 3x^2|_0^y I(y \in (0,1)) = 3y^2 I(0<y<1).$$

Please check that this is the pdf.

Then for $y \in (0,1)$

$$E(X|Y=y) = \frac{\int_0^y x6xdx}{3y^2} = \frac{2x^3|_0^y}{3y^2} = 2y^3/(3y^2) = (2/3)y.$$

4. Problem 14.7. Here

$$f^{X,Y}(x,y) = 2I(0<y<x<1).$$

Because now we have some experience in taking conditional expectations, I will write everything faster.

(a) For $x \in (0,1)$

$$E(Y|X=x) = \int_{-\infty}^{\infty} y f^{Y|X}(y|x)dy = \frac{\int_{-\infty}^{\infty} y f^{Y,X}(y,x)dy}{\int_{-\infty}^{\infty} f^{X,Y}(x,y)dy}$$

$$= \frac{2\int_0^x ydy}{2\int_0^x dy} = \frac{x^2}{2x} = x/2.$$

For $y \in (0,1)$

$$E(X|Y=y) = \frac{\int_{-\infty}^{\infty} x f^{Y,X}(y,x)dx}{\int_{-\infty}^{\infty} f^{X,Y}(x,y)dx}$$

$$= \frac{2\int_y^1 xdx}{2\int_y^1 dx} = \frac{1-y^2}{2(1-y)} = (1+y)/2.$$

In the last equality I used $1-y^2 = (1-y)(1+y)$.

(b)

$$E(X^mY^n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^m y^n f^{X,Y}(x,y)dxdy = \int_0^1 x^m \Big[\int_0^x y^n f^{X,Y}(x,y)dy\Big]dx$$

$$= 2\int_0^1 x^m[\int_0^x y^n dy]dx = 2\int_0^1 x^m (n+1)^{-1} x^{n+1} dx$$

$$= 2(n+1)^{-1}\int_0^1 x^{m+n+1}dx = \frac{2}{(n+1)(m+n+2)}.$$

5. Problem 14.8. Here

$$f^{X,Y}(x,y) = 24xyI(x>0)I(y>0)I(x+y<1).$$

Then for $x \in (0,1)$

$$E(Y|X=x) = \frac{\int_{-\infty}^{\infty} y f^{X,Y}(x,y)dy}{\int_{-\infty}^{\infty} f^{X,Y}(x,y)dy}$$

$$= \frac{\int_0^{1-x} y f^{X,Y}(x,y)dy}{\int_0^{1-x} f^{X,Y}(x,y)dy} = \frac{24x\int_0^{1-x} y^2 dy}{24x\int_0^{1-x} y dy}$$

$$= \frac{(1/3)(1-x)^3}{(1/2)(1-x)^2} = (2/3)(1-x).$$

6. Problem 14.9. Here

$$f^{X,Y}(x,y) = I(-y<x<y)I(0<y<1)$$

$$= I(y>|x|)I(0<y<1) = I(|x|<y<1).$$

[Am I correct with these indicator function manipulations?]

Then

$$Cov(X,Y) = E\{(X-E(X))(Y-E(Y))\} = E(XY) - E(X)E(Y).$$

Let us calculate the two terms.

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf^{XY}(x,y)dxdy = \int_{-1}^1 x[\int_{|x|}^1 ydy]dx$$

$$= \int_{-1}^1 x(1/2)(1-x^2)dx = 0.$$

The last equality holds because the integrable function is odd (or you can calculate the integral directly). Also,

$$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf^{XY}(x,y)dxdy = \int_{-1}^1 [\int_{|x|}^1 dy]dx = 0.$$

[The last equality holds by the same reason.] This yields $Cov(X,Y) = 0$, which in its turn implies $\rho = 0$.

At the same time, X and Y are dependent because you cannot write the joint density as a product $g_1(x)g_2(y)$. I finish here, but to be safe than sorry, it is a good idea to calculate two marginals and show that the joint pdf is not the product of the two marginal pdfs.

7. Problem 14.10. It is given that

$$E(Y|X=x) = a + bx, \quad Var(Y|X=x) = c,$$

where a, b, c are constants. [By the way, can c be negative?]

My guess is that the authors would like you to use Theorem 14.1 here, but I will indulge myself by using another approach. Write:

$$Cov(X,Y) = E\{(X-E(X))(Y-E(Y))\} = E\{E\{(X-E(X))(Y-E(Y))|X\}\}$$

$$= E\{(X-E(X))[E(Y|X)-E(Y)]\} = E\{(X-E(X))[a+bX-(a+bE(X))]\}$$

$$= bE\{(X-E(X))(X-E(X))\} = bVar(X). \qquad (1)$$

Using (1) we get

$$\rho = \frac{Cov(X,Y)}{[Var(X)Var(Y)]^{1/2}} = \frac{b[Var(X)]^{1/2}}{[Var(Y)]^{1/2}}.$$

Now I would like to remind you a classical formula (try to check it)

$$Var(Y) = Var(E(Y|X)) + E(Var(Y|X)). \qquad (2)$$

It yields for our case

$$Var(Y) = b^2 Var(X) + c. \qquad (3)$$

Combining (3),(1) and given $c = Var(Y|X)$, we get

$$Var(Y|X) = c = Var(Y) - b^2 Var(X)$$

$$= \frac{Var(Y) - b^2 Var(X)}{Var(Y)} Var(Y) = Var(Y)(1-\rho^2).$$

What was wished to verify..