

## 14.2: Problems on Conditional Expectation, Regression

### For the distributions in Exercises 1-3

- Determine the regression curve of  $Y$  on  $X$  and compare with the regression line of  $Y$  on  $X$ .
- For the function  $Z = g(X, Y)$  indicated in each case, determine the regression curve of  $Z$  on  $X$ .

#### Exercise 14.2.1

(See Exercise 17 from "[Problems on Mathematical Expectation](#)"). The pair  $\{X, Y\}$  has the joint distribution (in file [npr08\\_07.m](#)):

$$P(X = t, Y = u)$$

t =	-3.1	-0.5	1.2	2.4	3.7	4.9
u = 7.5	0.0090	0.0396	0.0594	0.0216	0.0440	0.0203
4.1	0.0495	0	0.1089	0.0528	0.0363	0.0231
-2.0	0.0405	0.1320	0.0891	0.0324	0.0297	0.0189
-3.8	0.0510	0.0484	0.0726	0.0132	0	0.0077

The regression line of  $Y$  on  $X$  is  $u = 0.5275t + 0.6924$

$$Z = X^2Y + |X + Y|$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = 0.5275t + 0.6924$

[npr08\\_07](#)

Data are in X, Y, P

jcalc

- - - - -

```
EYx = sum(u.*P)./sum(P);
```

```
disp([X;EYx]')
```

```
-3.1000 -0.0290
```

```
-0.5000 -0.6860
```

```
1.2000 1.3270
```

```
2.4000 2.1960
```

```
3.7000 3.8130
```

```
4.9000 2.5700
```

```
G = t.^2.*u + abs(t+u);
```

```
EZx = sum(G.*P)./sum(P);
```

```
disp([X;EZx]')
```

```
-3.1000 4.0383
```

```
-0.5000 3.5345
```

```
1.2000 6.0139
```

```
2.4000 17.5530
```

```
3.7000 59.7130
```

```
4.9000 69.1757
```

## Exercise 14.2.2

(See Exercise 18 from "[Problems on Mathematical Expectation](#)"). The pair  $\{X, Y\}$  has the joint distribution (in file [npr08\\_08.m](#)):

$$P(X = t, Y = u)$$

t =	1	3	5	7	9	11	13	15	17	19
u = 12	0.0156	0.0191	0.0081	0.0035	0.0091	0.0070	0.0098	0.0056	0.0091	0.0049
10	0.0064	0.0204	0.0108	0.0040	0.0054	0.0080	0.0112	0.0064	0.0104	0.0056
9	0.0196	0.0256	0.0126	0.0060	0.0156	0.0120	0.0168	0.0096	0.0056	0.0084
5	0.0112	0.0182	0.0108	0.0070	0.0182	0.0140	0.0196	0.0012	0.0182	0.0038
3	0.0060	0.0260	0.0162	0.0050	0.0160	0.0200	0.0280	0.0060	0.0160	0.0040
-1	0.0096	0.0056	0.0072	0.0060	0.0256	0.0120	0.0268	0.0096	0.0256	0.0084
-3	0.0044	0.0134	0.0180	0.0140	0.0234	0.0180	0.0252	0.0244	0.0234	0.0126
-5	0.0072	0.0017	0.0063	0.0045	0.0167	0.0090	0.0026	0.0172	0.0217	0.0223

The regression line of  $Y$  on  $X$  is  $u = -0.2584t + 5.6110$

$$Z = I_Q(X, Y)\sqrt{X}(Y - 4) + I_{Q^c}(X, Y)XY^2 \quad Q = \{(t, u) : u \leq t\}$$

### Answer

The regression line of  $Y$  on  $X$  is  $u = -0.2584t + 5.6110$

[npr08\\_08](#)

Data are in X, Y, P

jcalc

```
- - - - -
EYx = sum(u.*P)./sum(P);
disp([X;EYx]')
```

```
1.0000    5.5350
3.0000    5.9869
5.0000    3.6500
7.0000    2.3100
9.0000    2.0254
11.0000    2.9100
13.0000    3.1957
15.0000    0.9100
17.0000    1.5254
19.0000    0.9100
```

M = u<=t;

G = (u-4).\*sqrt(t).\*M + t.\*u.^2.\*(1-M);

EZx = sum(G.\*P)./sum(P);

```
disp([X;EZx]')
```

```
1.0000    58.3050
3.0000   166.7269
5.0000   175.9322
```

```

7.0000 185.7896
9.0000 119.7531
11.0000 105.4076
13.0000 -2.8999
15.0000 -11.9675
17.0000 -10.2031
19.0000 -13.4690

```

### Exercise 14.2.3

(See Exercise 19 from "[Problems on Mathematical Expectation](#)"). Data were kept on the effect of training time on the time to perform a job on a production line.  $X$  is the amount of training, in hours, and  $Y$  is the time to perform the task, in minutes. The data are as follows (in file [npr08\\_09.m](#)):

$$P(X = t, Y = u)$$

t =	1	1.5	2	2.5	3
u = 5	0.039	0.011	0.005	0.001	0.001
4	0.065	0.070	0.050	0.015	0.010
3	0.031	0.061	0.137	0.051	0.033
2	0.012	0.049	0.163	0.058	0.039
1	0.003	0.009	0.045	0.025	0.017

The regression line of  $Y$  on  $X$  is  $u = -0.7793t + 4.3051$ .

$$Z = (Y - 2.8)/X$$

### Answer

The regression line of  $Y$  on  $X$  is  $u = -0.7793t + 4.3051$ .

[npr08\\_09](#)

Data are in X, Y, P

```
jcalc
```

```
- - - - -
```

```
EYx = sum(u.*P)./sum(P);
```

```
disp([X;EYx]')
```

```
1.0000 3.8333
```

```
1.5000 3.1250
```

```
2.0000 2.5175
```

```
2.5000 2.3933
```

```
3.0000 2.3900
```

```
G = (u - 2.8)./t;
```

```
EZx = sum(G.*P)./sum(P);
```

```
disp([X;EZx]')
```

```
1.0000 1.0333
```

```
1.5000 0.2167
```

```
2.0000 -0.1412
```

2.5000	-0.1627
3.0000	-0.1367

### For the joint densities in Exercises 4-11 below

- Determine analytically the regression curve of  $Y$  on  $X$  and compare with the regression line of  $Y$  on  $X$ .
- Check these with a discrete approximation.

#### Exercise 14.2.4

(See Exercise 10 from "[Problems On Random Vectors and Joint Distributions](#)", Exercise 20 from "[Problems on Mathematical Expectation](#)", and Exercise 23 from "[Problems on Variance, Covariance, Linear Regression](#)").  $f_{XY}(t, u) = 1$  for  $0 \leq t \leq 1$ ,  $0 \leq u \leq 2(1-t)$ .

The regression line of  $Y$  on  $X$  is  $u = 1 - t$ .

$$f_X(t) = 2(1-t), 0 \leq t \leq 1$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = 1 - t$ .

$$f_{Y|X}(u|t) = \frac{1}{2(1-t)}, 0 \leq t \leq 1, 0 \leq u \leq 2(1-t)$$

$$E[Y|X=t] = \frac{1}{2(1-t)} \int_0^{2(1-t)} u du = 1 - t, 0 \leq t \leq 1$$

```
tuappr: [0 1] [0 2] 200 400 u<=2*(1-t)
- - - - -
EYx = sum(u.*P)./sum(P);
plot(X,EYx) % Straight line thru (0,1), (1,0)
```

#### Exercise 14.2.5

(See Exercise 13 from "[Problems On Random Vectors and Joint Distributions](#)", Exercise 23 from "[Problems on Mathematical Expectation](#)", and Exercise 24 from "[Problems on Variance, Covariance, Linear Regression](#)").  $f_{XY}(t, u) = \frac{1}{8}(t+u)$  for  $0 \leq t \leq 2, 0 \leq u \leq 2$ .

The regression line of  $Y$  on  $X$  is  $u = -t/11 + 35/33$ .

$$f_X(t) = \frac{1}{4}(t+1), 0 \leq t \leq 2$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = -t/11 + 35/33$ .

$$f_{Y|X}(u|t) = \frac{(t+u)}{2(t+1)}, 0 \leq t \leq 2, 0 \leq u \leq 2$$

$$E[Y|X=t] = \frac{1}{2(t+1)} \int_0^2 (tu + u^2) du = 1 + \frac{1}{3t+3}, 0 \leq t \leq 2$$

```
tuappr: [0 2] [0 2] 200 200 (1/8)*(t+u)
EYx = sum(u.*P)./sum(P);
eyx = 1 + 1./(3*X+3);
plot(X,EYx,X,eyx) % Plots nearly indistinguishable
```

### Exercise 14.2.6

(See Exercise 15 from "Problems On Random Vectors and Joint Distributions", Exercise 25 from "Problems on Mathematical Expectation", and Exercise 25 from "Problems on Variance, Covariance, Linear Regression").  $f_{XY}(t, u) = \frac{3}{88}(2t + 3u^2)$  for  $0 \leq t \leq 2, 0 \leq u \leq 1 + t$ .

The regression line of  $Y$  on  $X$  is  $u = 0.0958t + 1.4876$

$$f_X(t) = \frac{3}{88}(1+t)(1+4t+t^2) = \frac{3}{88}(1+5t+5t^2+t^3), \quad 0 \leq t \leq 2$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = 0.0958t + 1.4876$

$$\begin{aligned} f_{Y|X}(u|t) &= \frac{2t + 3u^2}{(1+t)(1+4t+t^2)} \quad 0 \leq u \leq 1+t \\ E[Y|X=t] &= \frac{1}{(1+t)(1+4t+t^2)} \int_0^{1+t} (2tu + 3u^3) du \\ &= \frac{(t+1)(t+3)(3t+1)}{4(1+4t+t^2)}, \quad 0 \leq t \leq 2 \end{aligned}$$

```
tuappr: [0 2] [0 3] 200 300 (3/88)*(2*t + 3*u.^2).*(u<=1+t)
EYx = sum(u.*P)./sum(P);
eyx = (X+1).*(X+3).*(3*X+1)./(4*(1 + 4*X + X.^2));
plot(X,EYx,X,eyx) % Plots nearly indistinguishable
```

### Exercise 14.2.7

(See Exercise 16 from "Problems On Random Vectors and Joint Distributions", Exercise 26 from "Problems on Mathematical Expectation", and Exercise 26 from "Problems on Variance, Covariance, Linear Regression").  $f_{XY}(t, u) = 12t^2u$  on the parallelogram with vertices

$$(-1, 0), (0, 0), (1, 1), (0, 1)$$

The regression line of  $Y$  on  $X$  is  $u = (4t + 5)/9$ .

$$f_X(t) = I_{[-1,0]}(t)6t^2(t+1)^2 + I_{(0,1]}(t)6t^2(1-t^2)$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = (23t + 4)/18$ .

$$\begin{aligned} f_{Y|X}(u|t) &= I_{[-1,0]}(t) \frac{2u}{(t+1)^2} + I_{(0,1]}(t) \frac{2u}{(1-t^2)} \quad \text{on the parallelogram} \\ E[Y|X=t] &= I_{[-1,0]}(t) \frac{1}{(t+1)^2} \int_0^{t+1} 2u du + I_{(0,1]}(t) \frac{1}{(1-t^2)} \int_t^1 2u du \\ &= I_{[-1,0]}(t) \frac{2}{3}(t+1) + I_{(0,1]}(t) \frac{2}{3} \frac{t^2 + t + 1}{t + 1} \end{aligned}$$

```
tuappr: [-1 1] [0 1] 200 100 12*t.^2.*u.*((u<= min(t+1,1))&(u>=max(0,t)))
EYx = sum(u.*P)./sum(P);
M = X<=0;
```

```
eyx = (2/3)*(X+1).*M + (2/3)*(1-M).*(X.^2 + X + 1)./(X + 1);
plot(X,EYx,X,eyx) % Plots quite close
```

### Exercise 14.2.8

(See Exercise 17 from "Problems On Random Vectors and Joint Distributions", Exercise 27 from "Problems on Mathematical Expectation", and Exercise 27 from "Problems on Variance, Covariance, Linear Regression").  $f_{XY}(t, u) = \frac{24}{11}tu$  for  $0 \leq t \leq 2, 0 \leq u \leq \min\{1, 2-t\}$ .

The regression line of  $Y$  on  $X$  is  $u = (-124t + 368)/431$

$$f_X(t) = I_{[0,1]}(t) \frac{12}{11}t + I_{(1,2]}(t) \frac{12}{11}t(2-t)^2$$

### Answer

The regression line of  $Y$  on  $X$  is  $u = (-124t + 368)/431$

$$\begin{aligned} f_{Y|X}(u|t) &= I_{[0,1]}(t)2u + I_{(1,2]}(t)\frac{2u}{(2-t)^2} \\ E[Y|X=t] &= I_{[0,1]}(t) \int_0^1 2u^2 du + I_{(1,2]}(t) \frac{1}{(2-t)^2} \int_0^{2-t} 2u^2 du \\ &= I_{[0,1]}(t) \frac{2}{3} + I_{(1,2]}(t) \frac{2}{3}(2-t) \end{aligned}$$

```
tuappr: [0 2] [0 1] 200 100 (24/11)*t.*u.*(u<=min(1,2-t))
EYx = sum(u.*P)./sum(P);
M = X <= 1;
eyx = (2/3)*M + (2/3).*(2 - X).*(1-M);
plot(X,EYx,X,eyx) % Plots quite close
```

### Exercise 14.2.9

(See Exercise 18 from "Problems On Random Vectors and Joint Distributions", Exercise 28 from "Problems on Mathematical Expectation", and Exercise 28 from "Problems on Variance, Covariance, Linear Regression").  $f_{XY}(t, u) = \frac{3}{23}(t+2u)$  for  $0 \leq t \leq 2, 0 \leq u \leq \max\{2-t, t\}$ .

The regression line of  $Y$  on  $X$  is  $u = 1.0561t - 0.2603$

$$f_X(t) = I_{[0,1]}(t) \frac{6}{23}(2-t) + I_{(1,2]}(t) \frac{6}{23}t^2$$

### Answer

The regression line of  $Y$  on  $X$  is  $u = 1.0561t - 0.2603$

$$\begin{aligned} f_{Y|X}(u|t) &= I_{[0,1]}(t) \frac{t+2u}{2(2-t)} + I_{(1,2]}(t) \frac{t+2u}{2t^2} \quad 0 \leq u \leq \max(2-t, t) \\ E[Y|X=t] &= I_{[0,1]}(t) \frac{1}{2(2-t)} \int_0^{2-t} (tu + 2u^2) du + I_{(1,2]}(t) \frac{1}{2t^2} \int_0^t (tu + 2u^2) du \\ &= I_{[0,1]}(t) \frac{1}{12}(t-2)(t-8) + I_{(1,2]}(t) \frac{7}{12}t \end{aligned}$$

```
tuappr: [0 2] [0 2] 200 200 (3/23)*(t+2*u).*(u<=max(2-t,t))
EYx = sum(u.*P)./sum(P);
M = X<=1;
eyx = (1/12)*(X-2).*(X-8).*M + (7/12)*X.*(1-M);
plot(X,EYx,X,eyx) % Plots quite close
```

### Exercise 14.2.10

(See Exercise 21 from "[Problems On Random Vectors and Joint Distributions](#)", Exercise 31 from "[Problems on Mathematical Expectation](#)", and Exercise 29 from "[Problems on Variance, Covariance, Linear Regression](#)").  $f_{XY}(t, u) = \frac{2}{13}(t + 2u)$  for  $0 \leq t \leq 2, 0 \leq u \leq \min\{2t, 3 - t\}$ .

The regression line of  $Y$  on  $X$  is  $u = -0.1359t + 1.0839$

$$f_X(t) = I_{[0,1]}(t) \frac{12}{13} t^2 + I_{(1,2]}(t) \frac{6}{13} (3 - t)$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = -0.1359t + 1.0839$

$$f_{Y|X}(t|u) = I_{[0,1]}(t) \frac{t+2u}{6t^2} + I_{(1,2]}(t) \frac{t+2u}{3(3-t)} \quad 0 \leq u \leq \max(2t, 3-t)$$

$$\begin{aligned} E[Y|X=t] &= I_{[0,1]}(t) \frac{1}{6t^2} \int_0^t (tu + 2u^2) du + I_{(1,2]}(t) \frac{1}{3(3-t)} \int_0^{3-t} (tu + 2u^2) du \\ &= I_{[0,1]}(t) \frac{11}{9} t + I_{(1,2]}(t) \frac{1}{18} (t^2 - 15t + 36) \end{aligned}$$

```
tuappr: [0 2] [0 2] 200 200 (2/13)*(t+2*u).*(u<=min(2*t,3-t))
EYx = sum(u.*P)./sum(P);
M = X<=1;
eyx = (11/9)*X.*M + (1/18)*(X.^2 - 15*X + 36).*(1-M);
plot(X,EYx,X,eyx) % Plots quite close
```

### Exercise 14.2.11

(See Exercise 22 from "[Problems On Random Vectors and Joint Distributions](#)", Exercise 32 from "[Problems on Mathematical Expectation](#)", and Exercise 30 from "[Problems on Variance, Covariance, Linear Regression](#)").

$f_{XY}(t, u) = I_{[0,1]}(t) \frac{3}{8} (t^2 + 2u) + I_{(1,2]}(t) \frac{9}{14} t^2 u^2$  for  $0 \leq u \leq 1$ .

The regression line of  $Y$  on  $X$  is  $u = 0.0817t + 0.5989$

$$f_X(t) = I_{[0,1]}(t) \frac{3}{8} (t^2 + 1) + I_{(1,2]}(t) \frac{3}{14} t^2$$

#### Answer

The regression line of  $Y$  on  $X$  is  $u = 0.0817t + 0.5989$

$$f_{Y|X}(t|u) = I_{[0,1]}(t) \frac{t^2 + 2u}{t^2 + 1} + I_{(1,2]}(t) 3u^2 \quad 0 \leq u \leq 1$$

$$E[Y|X=t] = I_{[0,1]}(t) \frac{1}{t^2 + 1} \int_0^1 (t^2 u + 2u^2) du + I_{(1,2]}(t) \int_0^1 3u^3 du$$

$$= I_{[0,1]}(t) \frac{3t^2 + 4}{6(t^2 + 1)} + I_{(1,2]}(t) \frac{3}{4}$$

```
tuappr: [0 2] [0 1] 200 100 (3/8)*(t.^2 + 2*u).*(t<=1) + ...
        (9/14)*t.^2.*u.^2.*(t>1)
EYx = sum(u.*P)./sum(P);
M = X<=1;
eyx = M.*(3*X.^2 + 4)./(6*(X.^2 + 1)) + (3/4)*(1 - M);
plot(X,EYx,X,eyx) % Plots quite close
```

### For the distributions in Exercises 12-16 below

- Determine analytically  $E[Z|X = t]$
- Use a discrete approximation to calculate the same functions.

#### Exercise 14.2.12

$f_{XY}(t, u) = \frac{3}{88}(2t + 3u^2)$  for  $0 \leq t \leq 2$ ,  $0 \leq u \leq 1 + t$ , (see Exercise 37 from "[Problems on Mathematical Expectation](#)", and [Exercise 14.2.6](#)).

$$f_X(t) = \frac{3}{88}(1+t)(1+4t+t^2) = \frac{3}{88}(1+5t+5t^2+t^3) \quad , 0 \leq t \leq 2$$

$$Z = I_{[0,1]}(X)4X + I_{(1,2]}(X)(X+Y)$$

#### Answer

$Z = I_M(X)4X + I_N(X)(X+Y)$ . Use of linearity, [\(CE8\)](#), and [\(CE10\)](#) gives

$$\begin{aligned} E[Z|X = t] &= I_M(t)4t + I_N(t)(t + E[Y|X = t]) \\ &= I_M(t)4t + I_N(t)\left(t + \frac{(t+1)(t+3)(3t+1)}{4(1+4t+t^2)}\right) \end{aligned}$$

% Continuation of [Exercise 14.2.6](#)

```
G = 4*t.*(t<=1) + (t + u).*(t>1);
EZx = sum(G.*P)./sum(P);
M = X<=1;
ezx = 4*X.*M + (X + (X+1).*(X+3).*(3*X+1)./(4*(1 + 4*X + X.^2))).*(1-M);
plot(X,EZx,X,ezx) % Plots nearly indistinguishable
```

#### Exercise 14.2.13

$f_{XY}(t, u) = \frac{24}{11}tu$  for  $0 \leq t \leq 2$ ,  $0 \leq u \leq \min\{1, 2-t\}$  (see Exercise 38 from "[Problems on Mathematical Expectation](#)", [Exercise 14.2.8](#)).

$$f_X(t) = I_{[0,1]}(t) \frac{12}{11}t + I_{(1,2]}(t) \frac{12}{11}t(2-t)^2$$

$$Z = I_M(X, Y) \frac{1}{2}X + I_M(X, Y)Y^2, \quad M = \{(t, u) : u > t\}$$

#### Answer

$$Z = I_M(X, Y) \frac{1}{2}X + I_M(X, Y)Y^2, \quad M = \{(t, u) : u > t\}$$



$$I_M(t, u) = I_{[0,1]}(t)I_{[t,1]}(u) \quad I_{M^c}(t, u) = I_{[0,1]}(t)I_{[0,t]}(u) + I_{(1,2]}(t)I_{[0,2-t]}(u)$$

$$E[Z|X=t] = I_{[0,1]}(t) \left[ \frac{t}{2} \int_t^1 2u \, du + \int_0^t u^2 \cdot 2u \, du \right] + I_{(1,2]}(t) \int_0^{2-t} u^2 \cdot \frac{2u}{(2-t)^2} \, du$$

$$= I_{[0,1]}(t) \frac{1}{2} t(1-t^2+t^3) + I_{(1,2]}(t) \frac{1}{2} (2-t)^2$$

% Continuation of Exercise 14.2.8

```
Q = u>t;
G = (1/2)*t.*Q + u.^2.*(1-Q);
EZx = sum(G.*P)./sum(P);
M = X <= 1;
ezx = (1/2)*X.*(1-X.^2+X.^3).*M + (1/2)*(2-X).^2.*(1-M);
plot(X,EZx,X,ezx) % Plots nearly indistinguishable
```

#### Exercise 14.2.14

$f_{XY}(t, u) = \frac{3}{23}(t+2u)$  for  $0 \leq t \leq 2$ ,  $0 \leq u \leq \max\{2-t, t\}$  (see Exercise 39 from "Problems on Mathematical Expectation", and Exercise 14.2.9).

$$f_X(t) = I_{[0,1]}(t) \frac{6}{23}(2-t) + I_{(1,2]}(t) \frac{6}{23}t^2$$

$$Z = I_M(X, Y)(X+Y) + I_{M^c}(X, Y)2Y, \quad M = \{(t, u) : \max(t, u) \leq 1\}$$

Answer

$$Z = I_M(X, Y)(X+Y) + I_{M^c}(X, Y)2Y, \quad M = \{(t, u) : \max(t, u) \leq 1\}$$

$$I_M(t, u) = I_{[0,1]}(t)I_{[0,1]}(u) \quad I_{M^c}(t, u) = I_{[0,1]}(t)I_{[1,2-t]}(u) + I_{(1,2]}(t)I_{[0,1]}(u)$$

$$E[Z|X=t] = I_{[0,1]}(t) \frac{1}{2(2-t)} \int_0^1 (t+u)(t+2u) \, du + \frac{1}{2-t} \int_1^{2-t} u(t+2u) \, du + I_{(1,2]}(t) 2E[Y|X=t]$$

$$= I_{[0,1]}(t) \frac{1}{12} \cdot \frac{2t^3 - 30t^2 + 69t - 60}{t-2} + I_{(1,2]}(t) \frac{7}{6} 2t$$

% Continuation of Exercise 14.2.9

```
M = X <= 1;
Q = (t<=1)&(u<=1);
G = (t+u).*Q + 2*u.*(1-Q);
EZx = sum(G.*P)./sum(P);
ezx = (1/12)*M.*(2*X.^3 - 30*X.^2 + 69*X - 60)./(X-2) + (7/6)*X.*(1-M);
plot(X,EZx,X,ezx)
```

#### Exercise 14.2.15

$f_{XY}(t, u) = \frac{2}{13}(t+2u)$ , for  $0 \leq t \leq 2$ ,  $0 \leq u \leq \min\{2t, 3-t\}$ . (see Exercise 31 from "Problems on Mathematical Expectation", and Exercise 14.2.10).

$$f_X(t) = I_{[0,1]}(t) \frac{12}{13}t^2 + I_{(1,2]}(t) \frac{6}{13}(3-t)$$

$$Z = I_M(X, Y)(X+Y) + I_{M^c}(X, Y)2Y^2, \quad M = \{(t, u) : t \leq 1, u \geq 1\}$$

# Answer

$$Z = I_M(X, Y)(X + Y) + I_{M^c}(X, Y)2Y^2, M = \{(t, u) : t \leq 1, u \geq 1\}$$

$$I_M(t, u) = I_{[0,1]}(t)I_{[1,2]}(u) \quad I_{M^c}(t, u) = I_{[0,1]}(t)I_{[0,1]}(u) + I_{(1,2]}(t)I_{[0,3-t]}(u)$$

$$E[Z|X=t] = I_{[0,1/2]}(t)\frac{1}{6t^2} \int_0^{2t} 2u^2(t+2u) du +$$

$$I_{(1/2,1]}(t)\left[\frac{1}{6t^2} \int_0^1 2u^2(t+2u) du + \frac{1}{6t^2} \int_1^{2t} (t+u)(t+2u) du\right] + I_{(1,2]}(t)\frac{1}{3(3-t)} \int_0^{3-t} 2u^2(t+2u) du$$

$$= I_{[0,1/2]}(t)\frac{32}{9}t^2 + I_{(1/2,1]}(t)\frac{1}{36} \cdot \frac{80t^3 - 6t^2 - 5t + 2}{t^2} + I_{(1,2]}(t)\frac{1}{9}(-t^3 + 15t^2 - 63t + 81)$$

```
tuappr: [0 2] [0 2] 200 200 (2/13)*(t + 2*u).*(u<=min(2*t,3-t))
M = (t<=1)&(u>=1);
Q = (t+u).*M + 2*(1-M).*u.^2;
EZx = sum(Q.*P)./sum(P);
N1 = X <= 1/2;
N2 = (X > 1/2)&(X<=1);
N3 = X > 1;
ezx = (32/9)*N1.*X.^2 + (1/36)*N2.*(80*X.^3 - 6*X.^2 - 5*X + 2)./X.^2 ...
      + (1/9)*N3.*(-X.^3 + 15*X.^2 - 63.*X + 81);
plot(X,EZx,X,ezx)
```

## Exercise 14.2.16

$f_{XY}(t, u) = I_{[0,1]}(t)\frac{3}{8}(t^2 + 2u) + I_{(1,2]}(t)\frac{9}{14}t^2u^2$ , for  $0 \leq u \leq 1$ . (see [Exercise 32](#) from "[Problems on Mathematical Expectation](#)", and [Exercise 14.2.11](#)).

$$f_X(t) = I_{[0,1]}(t)\frac{3}{8}(t^2 + 1) + I_{(1,2]}(t)\frac{3}{14}t^2$$

$$Z = I_M(X, Y)X + I_{M^c}(X, Y)XY, M = \{(t, u) : u \leq \min(1, 2 - t)\}$$

# Answer

$$Z = I_M(X, Y)X + I_{M^c}(X, Y)XY, M = \{(t, u) : u \leq \min(1, 2 - t)\}$$

$$E[X|X=t] = I_{[0,1]}(t) \int_0^1 \frac{t^3 + 2tu}{t^2 + 1} du + I_{(1,2]}(t) \left[ \int_0^{2-t} 3tu^2 du + \int_{2-t}^1 3tu^3 du \right]$$

$$= I_{[0,1]}(t)t + I_{(1,2]}(t)\left(-\frac{13}{4}t + 12t^2 - 12t^3 + 5t^4 - \frac{3}{4}t^5\right)$$

```
tuappr: [0 2] [0 1] 200 100 (t<=1).*(t.^2 + 2*u)./(t.^2 + 1) + 3*u.^2.*(t>1)
M = u<=min(1,2-t);
G = M.*t + (1-M).*t.*u;
EZx = sum(G.*P)./sum(P);
N = X<=1;
ezx = X.*N + (1-N).*(-(13/4)*X + 12*X.^2 - 12*X.^3 + 5*X.^4 - (3/4)*X.^5);
plot(X,EZx,X,ezx)
```

### Exercise 14.2.17

Suppose  $X \sim$  uniform on 0 through  $n$  and  $Y \sim$  conditionally uniform on 0 through  $i$ , given  $X = i$ .

- Determine  $E[Y]$  from  $E[Y|X = i]$ .
- Determine the joint distribution for  $\{X, Y\}$  for  $n = 50$  (see [Example 7](#) from "Conditional Expectation, Regression" for a possible approach). Use jcalc to determine  $E[Y]$ ; compare with the theoretical value.

#### Answer

- $E[Y|X = i] = i/2$ , so

$$E[Y] = \sum_{i=0}^n E[Y|X = i]P(X = i) = \frac{1}{n+1} \sum_{i=1}^n i/2 = n/4$$

- $P(X = i) = 1/(n+1)$ ,  $0 \leq i \leq n$ ,  $P(Y = k|X = i) = 1/(i+1)$ ,  $0 \leq k \leq i$ ; hence  $P(X = i, Y = k) = 1/(n+1)(i+1)$ ,  $0 \leq i \leq n$ ,  $0 \leq k \leq i$ .

```
n = 50; X = 0:n; Y = 0:n; P0 = zeros(n+1,n+1); for i = 0:n P0(i+1,1:i+1) =
(1/((n+1)*(i+1)))*ones(1,i+1); end P = rot90(P0); jcalc: X Y P - - - - -
- - EY = dot(Y,PY) EY = 12.5000 % Comparison with part (a): 50/4 = 12.5
```

### Exercise 14.2.18

Suppose  $X \sim$  uniform on 1 through  $n$  and  $Y \sim$  conditionally uniform on 1 through  $i$ , given  $X = i$ .

- Determine  $E[Y]$  from  $E[Y|X = i]$ .
- Determine the joint distribution for  $\{X, Y\}$  for  $n = 50$  (see [Example 7](#) from "Conditional Expectation, Regression" for a possible approach). Use jcalc to determine  $E[Y]$ ; compare with the theoretical value.

#### Answer

- $E[Y|X = i] = (i+1)/2$ , so

$$E[Y] = \sum_{i=1}^n E[Y|X = i]P(X = i) = \frac{1}{n+1} \sum_{i=1}^n \frac{i+1}{2} = \frac{n+3}{4}$$

- $P(X = i) = 1/n$ ,  $1 \leq i \leq n$ ,  $P(Y = k|X = i) = 1/i$ ,  $1 \leq k \leq i$ ; hence  $P(X = i, Y = k) = 1/ni$ ,  $1 \leq i \leq n$ ,  $1 \leq k \leq i$ .

```
n = 50; X = 1:n; Y = 1:n; P0 = zeros(n,n); for i = 1:n P0(i,1:i) =
(1/(n*i))*ones(1,i); end P = rot90(P0); jcalc: P X Y - - - - - EY =
dot(Y,PY) EY = 13.2500 % Comparison with part (a): 53/4 = 13.25
```

### Exercise 14.2.19

Suppose  $X \sim$  uniform on 1 through  $n$  and  $Y \sim$  conditionally binomial  $(i, p)$ , given  $X = i$ .

- Determine  $E[Y]$  from  $E[Y|X = k]$ .
- Determine the joint distribution for  $\{X, Y\}$  for  $n = 50$  and  $p = 0.3$ . Use jcalc to determine  $E[Y]$ ; compare with the theoretical value.

#### Answer

- $E[Y|X = i] = ip$ , so

$$E[Y] = \sum_{i=1}^n E[Y|X = i]P(X = i) = \frac{p}{n} \sum_{i=1}^n i = \frac{p(n+1)}{2}$$

- $P(X = i) = 1/n$ ,  $1 \leq i \leq n$ ,  $P(Y = k|X = i) = \text{ibinom}(i, p, 0 : i)$ ,  $0 \leq k \leq i$ .

```
n = 50; p = 0.3; X = 1:n; Y = 0:n; P0 = zeros(n,n+1); % Could use randbern for i
= 1:n P0(i,1:i+1) = (1/n)*ibinom(i,p,0:i); end P = rot90(P0); jcalc: X Y P - - -
- - - - - EY = dot(Y,PY) EY = 7.6500 % Comparison with part (a): 0.3*51/2 =
0.765
```

### Exercise 14.2.20

A number  $X$  is selected randomly from the integers 1 through 100. A pair of dice is thrown  $X$  times. Let  $Y$  be the number of sevens thrown on the  $X$  tosses. Determine the joint distribution for  $\{X, Y\}$  and then determine  $E[Y]$ .

#### Answer

a.  $P(X = i) = 1/n$ ,  $E[Y|X = i] = i/6$ , so

$$E[Y] = \frac{1}{6} \sum_{i=1}^n i/n = \frac{(n+1)}{12}$$

b.

```
n = 100; p = 1/6; X = 1:n; Y = 0:n; PX = (1/n)*ones(1,n); P0 = zeros(n,n+1); %
Could use randbern for i = 1:n P0(i,1:i+1) = (1/n)*ibinom(i,p,0:i); end P =
rot90(P0); jcalc EY = dot(Y,PY) EY = 8.4167 % Comparison with part (a): 101/12 =
8.4167
```

### Exercise 14.2.21

A number  $X$  is selected randomly from the integers 1 through 100. Each of two people draw  $X$  times, independently and randomly, a number from 1 to 10. Let  $Y$  be the number of matches (i.e., both draw ones, both draw twos, etc.). Determine the joint distribution and then determine  $E[Y]$ .

#### Answer

Same as [Exercise 14.2.20](#), except  $p = 1/10$ .  $E[Y] = (n+1)/20$

```
n = 100; p = 0.1; X = 1:n; Y = 0:n; PX = (1/n)*ones(1,n);
P0 = zeros(n,n+1); % Could use randbern
for i = 1:n
    P0(i,1:i+1) = (1/n)*ibinom(i,p,0:i);
end
P = rot90(P0);
jcalc
- - - - -
EY = dot(Y,PY)
EY = 5.0500 % Comparison with part (a): EY = 101/20 = 5.05
```

### Exercise 14.2.22

$E[Y|X = t] = 10t$  and  $X$  has density function  $f_X(t) = 4 - 2t$  for  $1 \leq t \leq 2$ . Determine  $E[Y]$ .

#### Answer

$$E[Y] = \int E[Y|X = t] f_X(t) dt = \int_1^2 10t(4 - 2t) dt = 40/3$$

### Exercise 14.2.23

$E[Y|X=t] = \frac{2}{3}(1-t)$  for  $0 \leq t < 1$  and  $X$  has density function  $f_X(t) = 30t^2(1-t)^2$  for  $0 \leq t \leq 1$ . Determine  $E[Y]$ .

**Answer**

$$E[Y] = \int E[Y|X=t]f_X(t) dt = \int_0^1 20t^2(1-t)^3 dt = 1/3$$

### Exercise 14.2.24

$E[Y|X=t] = \frac{2}{3}(2-t)$  and  $X$  has density function  $f_X(t) = \frac{15}{16}t^2(2-t)^2$   $0 \leq t < 2$ . Determine  $E[Y]$ .

**Answer**

$$E[Y] = \int E[Y|X=t]f_X(t) dt = \frac{5}{8} \int_0^2 t^2(2-t)^3 dt = 2/3$$

### Exercise 14.2.25

Suppose the pair  $\{X, Y\}$  is independent, with  $X \sim \text{Poisson}(\mu)$  and  $Y \sim \text{Poisson}(\lambda)$ . Show that  $X$  is conditionally binomial  $(n, \mu/(\mu+\lambda))$ , given  $X+Y=n$ . That is, show that

$$P(X=k|X+Y=n) = C(n, k)p^k(1-p)^{n-k}, \quad 0 \leq k \leq n, \text{ for } p = \mu/(\mu+\lambda)$$

**Answer**

$X \sim \text{Poisson}(\mu)$ ,  $Y \sim \text{Poisson}(\lambda)$ . Use of property (T1) and generating functions shows that  $X+Y \sim \text{Poisson}(\mu+\lambda)$

$$\begin{aligned} P(X=k|X+Y=n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} = \frac{P(X=k, Y=n-k)}{P(X+Y=n)} \\ &= \frac{e^{-\mu} \frac{\mu^k}{k!} e^{-\lambda} \frac{\lambda^{n-k}}{(n-k)!}}{e^{-(\mu+\lambda)} \frac{(\mu+\lambda)^n}{n!}} = \frac{n!}{k!(n-k)!} \frac{\mu^k \lambda^{n-k}}{(\mu+\lambda)^n} \end{aligned}$$

Put  $p = \mu/(\mu+\lambda)$  and  $q = 1-p = \lambda/(\mu+\lambda)$  to get the desired result.

### Exercise 14.2.26

Use the fact that  $g(X, Y) = g^*(X, Y, Z)$ , where  $g^*(t, u, v)$  does not vary with  $v$ . Extend property (CE10) to show

$$E[g(X, Y)|X=t, Z=v] = E[g(t, Y)|X=t, Z=v] \quad \text{a.s. } [P_{XZ}]$$

**Answer**

$$\begin{aligned} E[g(X, Y)|X=t, Z=v] &= E[g^*(X, Z, Y)|(X, Z)=(t, v)] = E[g^*(t, v, Y)|(X, Z)=(t, v)] \\ &= E[g(t, Y)|X=t, Z=v] \quad \text{a.s. } [P_{XZ}] \text{ by (CE10)} \end{aligned}$$

### Exercise 14.2.27

Use the result of Exercise 14.2.26 and properties (CE9a) and (CE10) to show that

$$E[g(X, Y)|Z=v] = \int E[g(t, Y)|X=t, Z=v]F_{X|Z}(dt|v) \quad \text{a.s. } [P_Z]$$

**Answer**

By (CE9),  $E[g(X, Y)|Z] = E\{E[g(X, Y)|X, Z]|Z\} = E[e(X, Z)|Z]$  a.s.

By (CE10),

$$E[e(X, Z)|Z = v] = E[e(X, v)|Z = v] = \int e(t, v)F_{X|Z}(dt|v) \text{ a.s.}$$

By Exercise 14.2.26,

$$\int E[g(X, Y)|X = t, Z = v]F_{X|Z}(dt|v) = \int E[g(t, Y)|X = t, Z = v]F_{X|Z}(dt|v) \text{ a.s. } [P_Z]$$

### Exercise 14.2.28

A shop which works past closing time to complete jobs on hand tends to speed up service on any job received during the last hour before closing. Suppose the arrival time of a job in hours before closing time is a random variable  $T \sim \text{uniform } [0, 1]$ . Service time  $Y$  for a unit received in that period is conditionally exponential  $\beta(2 - u)$ , given  $T = u$ . Determine the distribution uncton for  $Y$ .

**Answer**

$$F_Y(v) = \int F_{Y|T}(v|u)f_T(u) du = \int_0^1 (1 - e^{-\beta(2-u)v}) du = 1 - e^{-2\beta v} \frac{e^{\beta v} - 1}{\beta v} = 1 - e^{\beta v} \left[ \frac{1 - e^{-\beta v}}{\beta v} \right], 0 < v$$

### Exercise 14.2.29

Time to failure  $X$  of a manufactured unit has an exponential distribution. The parameter is dependent upon the manufacturing process. Suppose the parameter is the value of random variable  $H \sim \text{uniform on } [0.005, 0.01]$ , and  $X$  is conditionally exponential ( $u$ ), given  $H = u$ . Determine  $P(X > 150)$ . Determine  $E[X|H = u]$  and use this to determine  $E[X]$ .

**Answer**

$$F_{X|H}(t|u) = 1 - e^{-ut} \quad f_H(u) = \frac{1}{0.05} = 200, 0.005 \leq u \leq 0.01$$

$$F_X(t) = 1 - 200 \int_{0.005}^{0.01} e^{-ut} du = 1 - \frac{200}{t} [e^{-0.005t} - e^{-0.01t}]$$

$$P(X > 150) = \frac{200}{150} [e^{-0.75} - e^{-1.5}] \approx 0.3323$$

$$E[X|H = u] = 1/u \quad E[X] = 200 \int_{0.005}^{0.01} \frac{du}{u} = 200 \ln 2$$

### Exercise 14.2.30

A system has  $n$  components. Time to failure of the  $i$ th component is  $X_i$  and the class

$\{X_i : 1 \leq i \leq n\}$  is iid exponential ( $\lambda$ ). The system fails if any one or more of the components fails. Let  $W$  be the time to system failure. What is the probability the failure is due to the  $i$ th component?

*Suggestion.* Note that  $W = X_i$  iff  $X_j > X_i$ , for all  $j \neq i$ . Thus

$$\{W = X_i\} = \{(X_1, X_2, \dots, X_n) \in Q\}, Q = \{(t_1, t_2, \dots, t_n) : t_k > t_i, \forall k \neq i\}$$

$$P(W = X_i) = E[I_Q(X_1, X_2, \dots, X_n)] = E\{E[I_Q(X_1, X_2, \dots, X_n)|X_i]\}$$

**Answer**

Let  $Q = \{(t_1, t_2, \dots, t_n) : t_k > t_i, k \neq i\}$ . Then

$$P(W = X_i) = E[I_Q(X_1, X_2, \dots, X_n)] = E\{E[I_Q(X_1, X_2, \dots, X_n)|X_i]\}$$

$$= \int E[I_Q(X_1, X_2, \dots, t_i, \dots, X_n)] F_X(dt)$$

$$E[I_Q(X_1, X_2, \dots, t_i, \dots, X_n)] = \prod_{k \neq i} P(X_k > t) = [1 - F_X(t)]^{n-1}$$

If  $F_X$  is continuous, strictly increasing, zero for  $t < 0$ , put  $u = F_X(t)$ ,  $du = f_X(t) dt$ ,  $t = 0 \sim u = 0, t = \infty \sim u = 1$ .  
Then

$$P(W = X_i) = \int_0^1 (1-u)^{n-1} du = \int_0^1 u^{n-1} du = 1/n$$

This page titled [14.2: Problems on Conditional Expectation, Regression](#) is shared under a [CC BY 3.0](#) license and was authored, remixed, and/or curated by [Paul Pfeiffer](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.