CMPE 320: Probability, Statistics, and Random Processes

Lecture 10: Joint PMFs of multiple RVs

Spring 2018

Seung-Jun Kim

Multiple RVs

- Examples
  - Results of several tests in medical diagnosis
  - Workloads of several routers in computer networking
  - Results of multiple coin tosses

- In these examples, one needs consider events that involve multiple RVs simultaneously

Joint PMF

- Consider two RVs X and Y associated with the same experiment
- Joint PMF

$$p_{\underbrace{X,Y}_{\text{capitals RVs}}}(\overbrace{x,y}^{\text{lower-case}}) = P(\{X=x\}\cap\{Y=y\})$$
$$= P(X=x \text{ and } Y=y)$$
$$= P(X=x,\ Y=y) \quad \text{(simpler notation)}$$

Joint PMF as a table

- Joint PMF of 2 RVs can be arranged as a 2-D table
- Probability of set A of value pairs for (X,Y)

$$P((X,Y)\in A) = \sum_{(x,y)\in A} p_{X,Y}(x,y)$$

Joint PMF $p_{X,Y}(x,y)$ in tabular form

y \ x	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

ex) $A = \{(1,2),(2,2),(2,1)\}$

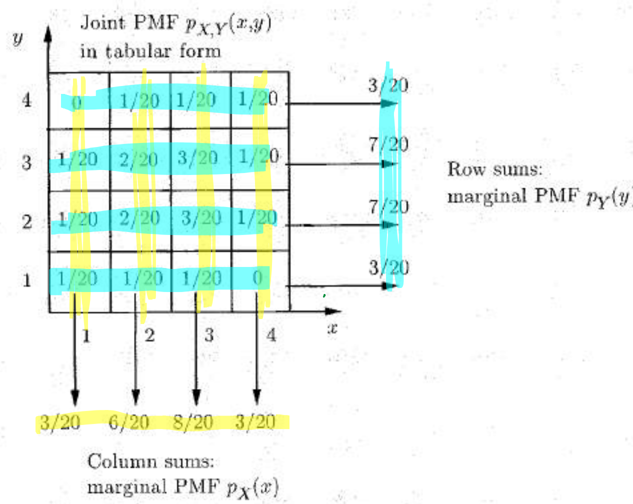
$$P((X,Y)\in A) = p_{X,Y}(1,2) + p_{X,Y}(2,2) + p_{X,Y}(2,1) = \frac{1}{20} + \frac{2}{20} + \frac{1}{20} = \frac{4}{20}$$

ex) Can you compute $P(X=1)$? Use the total probability

$$P(X=1) = P(X=1, Y=1) + P(X=1, Y=2) + P(X=1, Y=3) + P(X=1, Y=4)$$
$$= \sum_{\text{all } y} P(X=1, Y=y) = 0 + \frac{1}{20} + \frac{1}{20} + \frac{1}{20} = \frac{3}{20} = p_X(1)$$

Marginal PMF

- One can calculate PMFs of X and Y from the joint PMF



$$P_X(x) = \sum_{y} P_{X,Y}(x,y)$$

variable that is being marginalized

$$P_Y(y) = \sum_{x} P_{X,Y}(x,y)$$

Why? $P_X(x) = P(X=x)$

disjoint

$$\{X=x\} = \{X=x, Y=1\} \cup \{X=x, Y=2\} \cup \{X=x, Y=3\} \cup \{X=x, Y=4\}$$

$$P(X=x) = P(X=x, Y=1) + P(X=x, Y=2) + \dots + P(X=x, Y=4) = \sum_y \underbrace{P(X=x, Y=y)}_{P_{X,Y}(x,y)}$$

Functions of multiple RVs

- Consider a function Z = g(X,Y). What is the PMF of Z?

$$P_Z(z) = P(Z=z) = P(g(X,Y)=z)$$

$$= \sum_{(x,y):\, g(x,y)=z} P((X,Y)=(x,y))$$

sum over all (x, y) tuples which give rise to g(x, y) = z

$$\boxed{P_Z(z) = \sum_{(x,y):\, g(x,y)=z} P_{X,Y}(x,y)}$$

Expected values of Z = g(X,Y)

Recall for a single RV X : $E[g(X)] = \sum_x g(x) P_X(x)$

Likewise : $\boxed{E[g(X,Y)] = \sum_x \sum_y g(x,y) P_{X,Y}(x,y)}$

ex) $g(X,Y) = aX + bY + c$ (linear function of X, Y)

$E[aX+bY+c] = \sum_x \sum_y (ax+by+c) P_{X,Y}(x,y)$

$= \sum_x \sum_y ax P_{X,Y}(x,y) + \sum_x \sum_y by P_{X,Y}(x,y) + \sum_x \sum_y c P_{X,Y}(x,y)$

$= a\sum_x x \underbrace{\sum_y P_{X,Y}(x,y)}_{P_X(x)} + b\sum_y y \underbrace{\sum_x P_{X,Y}(x,y)}_{P_Y(y)} + c \underbrace{\sum_x \sum_y P_{X,Y}(x,y)}_{1}$

$= aE[X] + bE[Y] + c$

Example

- P_{XY} is given as in the table on the right. For Z = X + 2Y, compute E[Z] using 2 methods.
  1) Compute the PMF of Z first and then E[Z].
  2) Use E[aX+bY+c] = aE[X] + bE[Y] + c.

Joint PMF $p_{X,Y}(x,y)$ in tabular form

y \ x	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

2) $E[Z] = E[X+2Y] = E[X] + 2E[Y] = \frac{51}{20} + 2\cdot\frac{50}{20} = 7.55$

$P_X(x) = \begin{cases} \frac{3}{20}, & x=1 \\ \frac{6}{20}, & x=2 \\ \frac{8}{20}, & x=3 \\ \frac{3}{20}, & x=4 \end{cases}$

$E[X] = \sum_x x P_X(x)$

$= 1\cdot\frac{3}{20} + 2\cdot\frac{6}{20} + 3\cdot\frac{8}{20} + 4\cdot\frac{3}{20} = \frac{51}{20}$

$P_Y(y) = \begin{cases} \frac{3}{20}, & y=1 \\ \frac{7}{20}, & y=2,3 \\ \frac{3}{20}, & y=4 \end{cases}$

$E[Y] = \sum_y y P_Y(y)$

$= 1\cdot\frac{3}{20} + 2\cdot\frac{7}{20} + 3\cdot\frac{7}{20} + 4\cdot\frac{3}{20} = \frac{50}{20}$

1) $Z \in \{3, 4, \cdots, 12\}$

$P_Z(3) = P(Z=3) = P((X,Y)=(1,1)) = \frac{1}{20}$, $P_Z(4) = P(Z=4) = P((X,Y)=(2,1)) = \frac{1}{20}$

$P_Z(5) = P((X,Y) = (1,2), (3,1)) = \frac{1}{20} + \frac{1}{20} = \frac{2}{20}$

$P_Z(6) = P((X,Y) = (2,2), (4,1)) = \frac{2}{20} + 0 = \frac{2}{20}$

$\vdots$

$P_Z(12) = P((X,Y) = (4,4)) = \frac{1}{20}$

$E[Z] = \sum_z z \cdot P_Z(z) = 3 \cdot \frac{1}{20} + 4 \cdot \frac{1}{20} + 5 \cdot \frac{2}{20} + 6 \cdot \frac{2}{20} + \cdots + 12 \cdot \frac{1}{20}$

$= 7.55$

More than 2 RVs

Joint PMF of X, Y and Z

$P_{X,Y,Z}(x,y,z) = P(X=x, Y=y, Z=z)$

Marginalization

$P_{X,Y}(x,y) = \sum_z P_{X,Y,Z}(x,y,z)$

$P_X(x) = \sum_y \sum_z P_{X,Y,Z}(x,y,z)$

Expectation $E[aX + bY + cZ + d] = aE[X] + bE[Y] + cE[Z] + d$

Generalizes to n RVs

Example 2.10. Mean of the Binomial. Your probability class has 300 students and each student has probability 1/3 of getting an A, independent of any other student. What is the mean of X, the number of students that get an A?

$X_i = \begin{cases} 1 & \text{if the } i\text{-th student gets an A} \\ 0 & \text{otherwise} \end{cases}$, $i = 1, 2, \cdots, 300$

$X = X_1 + X_2 + \cdots + X_{300}$

$E[X] = E[X_1 + X_2 + \cdots + X_{300}] = E[X_1] + E[X_2] + \cdots + E[X_{300}]$

$p_{X_i}(x_i) = \begin{cases} \frac{1}{3} & \text{if } x_i = 1 \\ \frac{2}{3} & \text{if } x_i = 0 \end{cases}$ $\quad E[X_i] = 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}$

$i = 1, 2, \cdots, 300$

$E[X] = 300 \times \frac{1}{3} = 100.$

Example 2.11. The Hat Problem. Suppose that n people throw their hats in a box and then each picks one hat at random. (Each hat can be picked by only one person, and each assignment of hats to persons is equally likely.) What is the expected value of X, the number of people that get back their own hat?