Joint probability distributions

Let X and Y be two discrete rv's defined on the sample space $\mathcal{S}$.
The **joint probability mass function** $p(x, y)$ is
$p(x, y) = P(X = x, Y = y)$.
Note $p(x, y) \geq 0$ and $\sum_x \sum_y p(x, y) = 1$.
$P[(X, Y) \in A] = \sum \sum_{(x,y) \in A} p(x, y)$.

The **marginal probability mass function** of X and Y, denoted by $P_X(x)$ and $P_Y(y)$ are
$P_X(x) = \sum_y p(x, y), P_Y(y) = \sum_x P(x, y)$.

Example

Let X = number of meals in restaurants and Y = number of movies a randomly selected JMU student has on a typical weekend. The joint pmf of X and Y is given by

	p(x, y)	0	y 1	2	3
x	0	0.05	0.12	0.10	0.08
	1	0.06	0.12	0.11	0.09
	2	0.03	0.06	0.05	0.05
	3	0.03	0.02	0.02	0.01

$P(Y \geq 2) = p(0,2) + p(1,2) + p(2,2) + p(3,2) + p(0,3) + p(1,3) + p(2,3) + p(3,3) = 0.51$

The marginal distributions are:

$p_X(0) = 0.35, p_X(1) = 0.38, p_X(2) = 0.19, p_X(3) = 0.08$

$p_Y(0) = 0.17, p_Y(1) = 0.32, p_Y(2) = 0.28, p_Y(3) = 0.23.$

Exercise

Let X=number of heads and Y=number of heads minus the number of tails obtained in 3 flips of a balanced coin. Get the joint probability distribution of X and Y.

Joint pdf

Let X and Y be continuous rv's, then $f(x, y)$ is the **joint probability density function** for X and Y if for any two-dimensional set A
$P[(X, Y) \in A] = \int_A \int f(x, y)dxdy$. In particular,
$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y)dydx$.

The **marginal probability density function** of X and Y, denoted by $f_X(x)$ and $f_Y(y)$, are given by
$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$
and $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx$.

Example*

Given the joint pdf

$$f(x,y) = \begin{cases} \frac{2}{3}(x+2y), 0 < x < 1, 0 < y < 1 \\ 0, \text{otherwise} \end{cases}$$

$P(0 < X < 0.3, 0 < Y < 0.5) = \int_0^{0.3} \int_0^{0.5} \frac{2}{3}(x+2y)dydx =$
$\int_0^{0.3} (\frac{2}{3}xy + \frac{2}{3}y^2)|_{y=0}^{0.5} dx = \int_0^{0.3} (\frac{1}{3}x + \frac{1}{6})dx = \frac{x^2}{6} + \frac{x}{6}|_0^{0.3} = 0.02$

$f_X(x) = \int_0^1 \frac{2}{3}(x+2y)dy = (\frac{2}{3}xy + \frac{2}{3}y^2)|_{y=0}^1 = \frac{2}{3}(x+1), 0 < x < 1$
$f_Y(y) = \int_0^1 \frac{2}{3}(x+2y)dx = \frac{1}{3}(1+4y), 0 < y < 1.$

exercise

Given the joint pdf

$$f(x,y) = \begin{cases} \frac{3}{5}x(y+x), 0 < x < 1, 0 < y < 2 \\ 0, \text{otherwise} \end{cases}$$

Find $P(0 < X < \frac{1}{2}, 0 < Y < 1)$.
Find $f_Y(y)$.

Independent rv's

Two rv's X and Y are said to be **independent** if for every pair of (x, y),
$p(x, y) = P_X(x)P_Y(y)$, when X and Y are discrete
or $f(x, y) = f_X(x)f_Y(y)$ when X and Y are continuous.

example: meals and movie example:
$p(0, 0) = 0.05 \neq p_X(0) * p_Y(0) = 0.35 * 0.17$.

example *: $f_X(x)f_Y(y) = \frac{2}{3}(x + 1)\frac{1}{3}(1 + 4y) \neq f(x, y)$, so X and Y are dependent.

Expected values, covariance and correlation

Let X and Y be two rv's with joint pmf $p(x, y)$ or $f(x, y)$, then

$$E(h(X, Y)) = \begin{cases} \sum_x \sum_y h(x, y)p(x, y), \text{if X and Y are discrete} \\ \int \int h(x, y)f(x, y)dxdy, \text{if X and Y are continuous} \end{cases}$$

The **covariance** between X and Y is
$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$

$$= \begin{cases} \sum_x \sum_y (x - \mu_X)(y - \mu_Y)p(x, y), \text{X, Y discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)f(x, y)dxdy, \text{X, Y continuous} \end{cases}$$

proposition: $Cov(X, Y) = E(XY) - \mu_X \mu_Y$.

example

Meals and movies example:

			y			
	p(x, y)	0	1	2	3	p(x)
x	0	0.05	0.12	0.10	0.08	**0.35**
	1	0.06	0.12	0.11	0.09	**0.38**
	2	0.03	0.06	0.05	0.05	**0.19**
	3	0.03	0.02	0.02	0.01	**0.08**
	p(y)	**0.17**	**0.32**	**0.28**	**0.23**	

Note $\mu_X = \sum xp(x) = 1, \mu_Y = \sum yp(y) = 1.57$,
$E(XY) = \sum_x \sum_y xyp(x, y) =$
$1 * 1 * 0.12 + 1 * 2 * 0.11 + 1 * 3 * 0.09 + 2 * 1 * 0.06 + 2 * 2 * 0.05 + 2 * 3 * 0.05 + 3 * 1 * 0.02 + 3 * 2 * 0.02 + 3 * 3 * 0.01 = 1.50$,
so $Cov(X, Y) = E(XY) - \mu_X \mu_Y = 1.50 - 1 * 1.57 = -0.07$

The **correlation coefficient** of X an Y, denoted by corr(X, Y), or $\rho_{X,Y}$ or just ρ, is
$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$.
Meals and movies example:
$E(X^2) = 1^2 * 0.38 + 2^2 * 0.19 + 3^2 * 0.08 = 1.86, \sigma_X^2 = 1.86 - 1^2 = 0.86, \sigma_X = 0.927$, and $\sigma_Y = 1.022$,
so $\rho_{X,Y} = \frac{-0.07}{0.927*1.022} = -0.07$
proposition:
1. If a and c are both positive or negative, $Corr(aX + b, cY + d) = Corr(X, Y)$.
2. $-1 \leq \rho \leq 1$.
3. If X and Y are independent, then $\rho = 0$ but $\rho = 0$ does not necessarily imply independence.
4. $\rho = \pm 1$ iff $Y = aX + b$ for $a \neq 0$.

Suppose X and Y are independent with pdf's
$f_X(x) = 3x^2, 0 \leq x \leq 1$
$f_Y(y) = 2y, 0 \leq y \leq 1.$
Find $E(\frac{X}{Y})$.
Optional: Find $P(X \geq Y)$.

The joint pdf of X and Y is
$f(x, y) = 6x^2y, 0 \leq x \leq 1, 0 \leq y \leq 1.$
$E(\frac{X}{Y}) = \int_0^1 \int_0^1 (\frac{x}{y}) 6x^2 y dxdy = 1.5.$
$P(X \geq Y) = \int_0^1 \int_0^x 6x^2 y dydx = 0.6$

Conditional distribution:
Let X and Y be discrete r.v.s with joint pmf $p(x, y)$ and marginal pmf $p_X(x)$ and $p_Y(y)$. The conditional probability mass function of Y given $X = x$ is
$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$.
Let X(deductible on a car policy) and Y (deductible on a home policy) have the joint pmf below:

		y		
p(x,y)		0	100	200
x	100	0.20	0.10	0.20
	250	0.05	0.15	0.30

Find the conditional distribution of Y given $X = 250$.
$P(X = 250) = 0.5$,
$p_{Y|X}(0|250) = \frac{0.05}{0.50} = 0.1$,
$p_{Y|X}(100|250) = \frac{0.15}{0.50} = 0.3$
$p_{Y|X}(200|250) = \frac{0.30}{0.50} = 0.60$.

Conditional pdf

Let X and Y be two continuous r.v.s with joint pdf $f(x,y)$ and marginal pdf $f_X(x)$ and $f_Y(y)$ respectively. The conditional pdf of Y given $X = x$ is
$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$.

Example

Given the joint pdf

$$f(x,y) = \begin{cases} 24xy, 0 \le x \le 1, 0 \le y \le 1, x+y \le 1 \\ 0, \text{otherwise} \end{cases}$$

$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^{1-x} 24xydy = 12x(1-x)^2, 0 \le x \le 1.$
Then $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{24xy}{12x(1-x)^2}, 0 \le x \le 1, 0 \le y \le 1-x.$
Hence $f_{Y|X}(y|0.5) = \frac{2y}{(1-0.5)^2} = 8y, 0 \le y \le 0.5.$