ISyE 3770
Chapter 5: Joint Probability Distributions

Instructor: Dan Li

School of Industrial and Systems Engineering
Georgia Tech

Spring 2021

1. Joint Distributions
   - Marginal Distribution
   - Conditional Distribution
   - Independence
   - Covariance & Correlation

2. Linear Functions of Random Variables
   - Linear Function
   - Mean & Variance of a Linear Function
   - Reproductive Property of the Normal Distribution

Joint PMF for Discrete RV

- The joint probability mass function of the discrete random variables X and Y denoted as $f_{XY}(xy)$, satisfies the following conditions:

$$f_{XY}(xy) \geq 0$$

$$\sum_x \sum_y f_{XY}(xy) = 1$$

$$f_{XY}(xy) = \mathbb{P}(X = x, Y = y)$$

Example: Mobile Response Time

- The response time is the speed of page downloads and it is critical for a mobile Web site. As the response time increases, customers become more frustrated and potentially abandon the site for a competitive one. Let X denote the number of bars of service, and let Y denote the response time (to the nearest second) for a particular user and site. The joint probability distribution of X and Y is given below:

	***x* = Number of Bars of Signal Strength**		
***y* = Response time (nearest second)**	**1**	**2**	**3**
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

Example: Mobile Response Time

- Consider the joint pmf of X and Y in the example.

	x = Number of Bars of Signal Strength		
y = Response time (nearest second)	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

- Calculate $P(X \leq 2, Y < 3)$.
- Calculate $P(Y < 3)$.

Example: Mobile Response Time

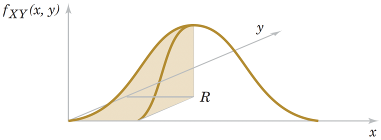
Joint PDF for Continuous RV

- The joint probability density function of the continuous random variables X and Y denoted as $f_{XY}(xy)$, satisfies the following conditions:

$$f_{XY}(xy) \geq 0$$

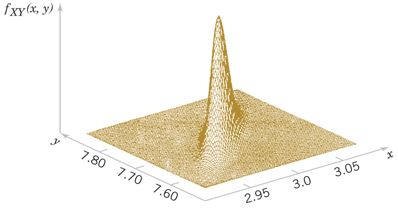
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(xy)\mathrm{d}x\mathrm{d}y = 1$$

$$\mathbb{P}((X,Y) \in R) = \int\int_{R} f_{XY}(xy)\mathrm{d}x\mathrm{d}y$$



Example: Injection-Molded Part

- Let the continuous r.v. X denote the length of one dimension of an injection-molded part, and the continuous r.v. Y denote the length of another dimension. The specifications for X and Y are (2.95 to 3.05) and (7.60 to 7.80) millimeters, respectively. What is the probability that the part satisfies both specifications?



Marginal Distribution: Discrete RV

- The marginal probability distribution for X is found by summing the probabilities in each column whereas the marginal probability distribution for Y is found by summing the probabilities in each row.

$$f_X(x) = \sum_y f(xy)$$

$$f_Y(y) = \sum_x f(xy)$$

y = Response time(nearest second)	x = Number of Bars of Signal Strength			
	1	2	3	$f(y)$
1	0.01	0.02	0.25	0.28
2	0.02	0.03	0.20	0.25
3	0.02	0.10	0.05	0.17
4	0.15	0.10	0.05	0.30
$f(x)$	0.20	0.25	0.55	1.00

Marginal probability distributions of X and Y

- How about for continuous random variables?

Marginal Distribution: Continuous RV

- If the joint probability density function of random variables X and Y is $f_{XY}(x,y)$, the marginal probability density functions of X and Y are:

$$f_X(x) = \int f_{XY}(x,y)dy$$

$$f_Y(y) = \int f_{XY}(x,y)dx$$

where the first integral is over all points in the range of (X,Y) for which $X = x$ and the second integral is over all points in the range of (X,Y) for which $Y = y$.

Mean & Variance of a Marginal Distribution

- $\mathbb{E}(X)$ and $V(X)$ can be obtained by first calculating the marginal probability distribution of X and then determining $\mathbb{E}(X)$ and $V(X)$ by the usual method.
- For discrete r.v. X:

$$E(X) = \sum_{X} x \cdot f_X(x) = \sum_{x} [x \cdot \sum_{y} f_{XY}(x, y)]$$

$$V(X) = \sum_{X} x^2 \cdot f_X(x) - \mu_X^2$$

Mean & Variance of a Marginal Distribution

- If the joint probability density function of random variables X and Y is $f_{XY}(x, y)$,
- For continuous r.v. X:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \cdot f_{XY}(x)dydx$$

$$V(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f_{XY}(x)dydx$$

Example: Mobile Response Time

- Consider the joint pmf of X and Y in the example.

	x = Number of Bars of Signal Strength		
y = Response time (nearest second)	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

- Calculate $E[X]$ and $V(X)$.
- Find the marginal probability distribution of X.

Example: Mobile Response Time

Conditional Distributions

- Given continuous random variables X and Y with joint probability density function $F_{XY}(x, y)$, the conditional probability density function of Y given $X = x$ is

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \text{for} \quad f_X(x) > 0$$

which satisfies the following properties:

1. $f_{Y|x}(y) \geq 0$
2. $\int f_{Y|x}(y)\mathrm{d}y = 1$
3. $\mathbb{P}(Y \in B|X = x) = \int_B f_{Y|x}(y)\mathrm{d}y$ for any set B in the range of Y

Conditional Mean & Variance

- The conditional mean of Y given $X = x$, denoted as $\mathbb{E}(Y|x)$ or $\mu_{Y|x}$ is:

$$\mathbb{E}[Y|x] = \int_y y \cdot f_{Y|x}(y)\mathrm{d}y$$

- The conditional variance of Y given $X = x$, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$ is:

$$V(Y|x) = \int_y (y - \mu_{Y|x})^2 \cdot f_{Y|x}(y)\mathrm{d}y = \int_y y^2 \cdot f_{Y|x}(y)\mathrm{d}y - \mu^2_{Y|x}$$

Example: Mobile Response Time

- Consider the joint pmf of X and Y in the example.

	x = Number of Bars of Signal Strength		
y = Response time (nearest second)	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

- Find the conditional distribution of Y given that $X = 1$.
- Calculate $E(Y|X = 1)$.

Example: Mobile Response Time

Independence

- For random variables X and Y, if any one of the following properties is true, the others are also true. Then X and Y are independent.
  1. $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$ for all x and y
  2. $f_{Y|x}(y) = f_Y(y)$ for all x and y with $f_X(x) > 0$
  3. $f_{X|y}(x) = f_X(x)$ for all x and y with $f_Y(y) > 0$
  4. $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B)$ for any set A and B in the range of X and Y, respectively.

Example: Mobile Response Time

- Consider the joint pmf of X and Y in the example.

	x = Number of Bars of Signal Strength		
y = Response time (nearest second)	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

- Are X and Y independent? Why?

Covariance & Correlation, and Independence

Given the joint distributions of X and Y, we define their covariance:

$$cov(X,Y) = \mathbb{E}\left[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

and we define their **correlation**:

$$cor(X,Y) = \frac{cov(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}.$$

X and Y are **uncorrelated** if $cov(X,Y) = 0$ or $cor(X,Y) = 0$.

Example

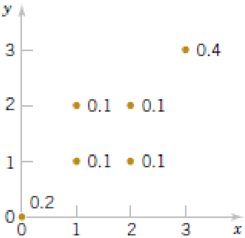


Figure 5-13 Discrete joint distribution, $f(x, y)$.

	x	y	f(x, y)	x-μX	y-μY	Prod
Joint	0	0	0.2	-1.8	-1.2	0.42
	1	1	0.1	-0.8	-0.2	0.01
	1	2	0.1	-0.8	0.8	-0.07
	2	1	0.1	0.2	-0.2	0.00
	2	2	0.1	0.2	0.8	0.02
	3	3	0.4	1.2	1.8	0.88
Marginal	0		0.2	covariance =		1.260
	1		0.2	correlation =		0.926
	2		0.2			
	3		0.4	Note the strong		
		0	0.2	positive correlation.		
		1	0.2			
		2	0.2			
		3	0.4			
Mean		μ_X =	1.8			
		μ_Y =	1.8			
StDev		σ_X =	1.1662			
		σ_Y =	1.1662			

Covariance, Correlation, and Independence

- If X and Y are independent, they are uncorrelated.
- If X and Y are uncorrelated, it doesn't necessarily mean they are independent!
- Special case: for **normal distributions**, if X and Y are uncorrelated, then they are independent. i.e., for normal distributions, independence $\Leftrightarrow$ uncorrelatedness.

Example

- Consider when X follows a Uniform distribution on $[-1, 1]$, and $Y = \begin{cases} X & X > 0 \\ -X & X \leq 0 \end{cases}$

Linear Function

- Given random variables X_1, X_2,...,X_p and constants $c_0, c_1, c_2, ..., c_p$,

$$Y = c_0 + c_1X_1 + C_2X_2 + ... + c_pX_p$$

is a **linear function** of $X_1, X_2, ..., X_p$

Mean & Variance of a Linear Function

- If $Y = c_0 + c_1X_1 + c_2X_2 + ... + c_pX_p$,

$$E(Y) = c_0 + c_1E(X_1) + c_2E(X_2) + ... + c_pE(X_p).$$

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + ... + c_p^2V(X_p) + 2\sum_i\sum_{j>i} c_ic_j cov(X_i, X_j)$$

- If $X_1, X_2, ..., X_p$ are independent,

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + ... + c_p^2V(X_p)$$

Mean & Variance of an Average

- If $\bar{X} = (X_1 + X_2 + ... + X_p)/p$ with $E(X_i) = \mu$ for $i = 1, 2, ..., p$,
$$E(\bar{X}) = \mu.$$
- If $X_1, X_2, ..., X_p$ are independent with $V(X_i) = \sigma^2$ for $i = 1, 2, ..., p$,
$$V(\bar{X}) = \sigma^2/p$$

Reproductive Property of the Normal Distribution

- If $X_1, X_2, ..., X_p$ are independent, normal random variables with $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$ for $i = 1, 2, ..., p$, and $Y = c_0 + c_1X_1 + c_2X_2 + ... + c_pX_p$, then Y is a **normal random variable** with

$$E(Y) = c_0 + c_1\mu_1 + c_2\mu_2 + ... + c_p\mu_p.$$

$$V(Y) = c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + ... + c_p^2\sigma_p^2$$

- In short: a linear function of normal random variables is still a normal random variable.