Bivariate and Multivariate Probability Distributions

- Often we are interested in more than one aspect of an experiment/trial.
- Will have more than one random variable.
- Interest in the probability of a combination of events (results of different aspects of the experiment).

Examples include:

- Price of crude oil (per barrel) and price per gallon of unleaded gasoline at your local station (per gallon).
- Level of different contaminants in soil samples.
- Probability of obtaining a certain sample mean and sample variance in a sample from a population.

Discrete Bivariate Distributions

- If X_1 and X_2 are discrete random variables, the joint distribution of X_1 and X_2 is given by the joint p.m.f. $p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$.
- This joint p.m.f. must satisfy
  - $p(x_1, x_2) \geq 0$ for all x_1 and x_2,
  - $\sum_{x_1} \sum_{x_2} p(x_1, x_2) = 1$.

Joint Distribution Function (Joint c.d.f.)

- The joint c.d.f. of two random variables X_1 and X_2 is given by $F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$ for $-\infty < x_1 < \infty$ and $-\infty < x_2 < \infty$.
- For discrete random variables X_1 and X_2 with joint p.m.f. $p(x_1, x_2)$, the joint c.d.f. is
  $F(x_1, x_2) = \sum_{y_1 \leq x_1} \sum_{y_2 \leq x_2} p(y_1, y_2)$

Properties of Joint c.d.f.

- $F(-\infty, -\infty) = F(x_1, -\infty) = F(-\infty, x_2) = 0$ for all x_1 and x_2; $F(\infty, \infty) = 1$.
- $F(y_1, y_2) - F(y_1, x_2) - F(x_1, y_2) + F(x_1, x_2) \geq 0$ if $y_1 > x_1$ and $y_2 > x_2$.

A Discrete Bivariate Example

Consider two balanced dice of which the first die has 3 faces marked "1" and other 3 faces marked "2" and the second die has 2 "1" faces and 2 "2" faces, and 2 "3" faces. Each die is rolled once.

X: number of "2"'s rolled.

Y: sum of the numbers on the top faces.

Find $p(x, y) = P(X = x, Y = y)$. Also find the c.d.f.

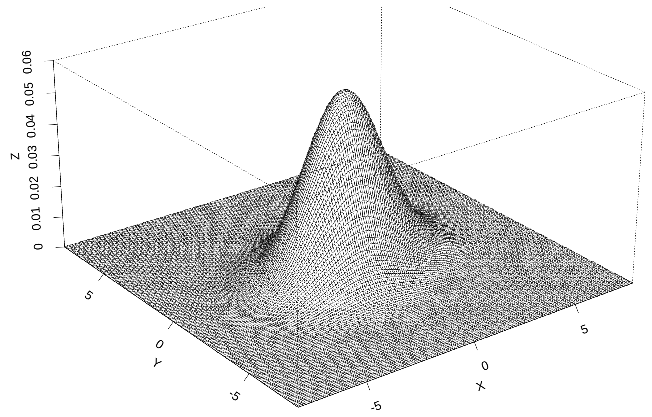
Bivariate and Continuous

- Random variables X_1 and X_2 are jointly continuous if their joint c.d.f $F(x_1, x_2)$ is continuous in both arguments.
- If X_1 and X_2 are jointly continuous, they have a joint density function or joint p.d.f.
- The joint density of X_1 and X_2 is $f(x_1, x_2)$ if for all $-\infty < x_1 < \infty$ and $-\infty < x_2 < \infty$

  1. $f(x_1, x_2) \geq 0$ and
  2. $F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(t_1, t_2) dt_2 dt_1$.

- Volume under the surface must be 1: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 dy_2 = 1$.

- Some calculations will require multiple integrals.



Bivariate normal density

Marginal Distributions

- If we're given the joint distribution for 2 or more variables, how can we find the distribution for just one of them?
- Discrete case:
  - Suppose X and Y are two discrete random variables with joint p.m.f. $p(x, y)$. We want the p.m.f. $p_1(x)$ of X and the p.m.f. $p_2(y)$ of Y.
  - $(X = x)$ can be written as a countable union of mutually exclusive events where the union is taken over all possible values of Y:

$$(X = x) = \cup_y (X = x, Y = y).$$

  - Since they're mutually exclusive, we sum the probabilities for all the different possible values of Y that can occur with $X = x$.
  - This leads to $p_1(x) = \sum_y p(x, y)$ and $p_2(y) = \sum_x p(x, y)$. These p_1 and p_2 are called the marginal p.m.f. of X and Y, respectively.
- Continuous case: If X and Y have a joint p.d.f. $f(x, y)$, then the marginal p.d.f. $f_1(x)$ of X and $f_2(y)$ of Y are given by $f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$ and $f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$.
- The marginal c.d.f. $F_1(x)$ and $F_2(y)$ of X and Y are given by $F_1(x) = F(x, \infty)$ and $F_2(y) = F(\infty, y)$.

Dice Example Continued

What are the marginal distributions for X and Y in our earlier dice example?

X: number of "2"'s rolled.

Y: sum of the numbers on the top faces.

	Y			
X	2	3	4	5
0	1/6	0	1/6	0
1	0	1/3	0	1/6
2	0	0	1/6	0

Conditional Distributions

Suppose we have two random variables X and Y with a joint p.m.f. or p.d.f. and we want to know the p.m.f. or p.d.f. of one given the value of the other.

- Discrete case: Use the definition of conditional probability $P(A|B) = \frac{P(A \cap B)}{P(B)}$. If the joint p.m.f. of X and Y is $p(x, y)$, and the marginal p.m.f. of Y is $p_2(y)$, then the conditional p.m.f. of X given $Y = y$ is $P(X = x|Y = y) = \frac{P(X=x,Y=y)}{P(Y=y)} = \frac{p(x,y)}{p_2(y)}$.
- Continuous case: If the joint p.d.f. of X and Y is $f(x, y)$ and the marginal p.d.f. of Y is $f_2(y)$, then the conditional p.d.f. of X given $Y = y$ is $f(x|y) = \frac{f(x,y)}{f_2(y)}$.

Dice Example Revisited

X: number of "2"'s rolled.

Y: sum of the numbers on the top faces.

	Y			
X	2	3	4	5
0	1/6	0	1/6	0
1	0	1/3	0	1/6
2	0	0	1/6	0

Give the conditional probability distribution of Y given X and X given Y in our dice example.

Independence of Random Variables

In chapter 2, we discussed independence and dependence of events A and B.

- A and B are independent if $P(A|B) = P(A)$ or $P(B|A) = P(B)$ or $P(A \cap B) = P(A)P(B)$.
- Otherwise, knowing A happened gives info about P(B) (and vice-versa) and A and B are dependent.

Extend this principle to random variables and their probability distributions/densities.

- Suppose the joint c.d.f. of X and Y is $F(x, y)$ and the marginals are $F_1(x)$ and $F_2(y)$, respectively. Then X and Y are independent if $F(x, y) = F_1(x)F_2(y)$ for all x and y.

- In discrete case, independence is equivalent to the condition: $p(x,y) = p_1(x)p_2(y)$ for all x and y.
- In continuous case, independence is equivalent to the condition: $f(x,y) = f_1(x)f_2(y)$ for all x and y.

Independence of Functions of Random Variables

If X and Y are independent and f is a function of only X and g is a function of only Y, then $f(X)$ and $g(Y)$ are independent.

Independence and Our Dice Example

X: number of "2"'s rolled.

Y: sum of the numbers on the top faces.

	Y			
X	2	3	4	5
0	1/6	0	1/6	0
1	0	1/3	0	1/6
2	0	0	1/6	0

In our dice example, were X and Y independent?

Expected Value

We just extend our ideas about expected value to more than one variable. If $g(X_1, X_2, ..., X_n)$ is a function of random variables $X_1, X_2, ..., X_n$ for which we are interested in finding the expected value:

Discrete case:

$$E[g(X_1, X_2, ..., X_n)]$$
$$= \sum_{x_1}\sum_{x_2}...\sum_{x_n} g(x_1, x_2, ..., x_n)p(x_1, x_2, ..., x_n)$$

Continuous case:

$$E[g(X_1, X_2, ..., X_n)]$$
$$= \int_{x_1}\int_{x_2}...\int_{x_n} g(x_1, x_2, ..., x_n)f(x_1, x_2, ..., x_n)dx_n...dx_2dx_1$$

Properties of Expected Value

Also, we still have the same properties for expected values that we discussed before:

- $E(c) = c$ where c is a constant.
- $E[cg(X_1, X_2, ..., X_n)] = cE[g(X_1, X_2, ..., X_n)]$.
- $E\left[\sum_{i=1}^{k} g_i(X_1, X_2, ..., X_n)\right] = \sum_{i=1}^{k} E[g_i(X_1, X_2, ..., X_n)]$.

Independence and Expectation

- If X and Y are independent random variables, then $E(XY) = E(X)E(Y)$.
- The converse is not true, i.e., there are dependent random variables X and Y for which $E(XY) = E(X)E(Y)$.

Covariance

- For two random variables X and Y the covariance is defined as: $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$, where $E(X) = \mu_X$ and $E(Y) = \mu_Y$.
- $Cov(X,X) = V(X)$.
- The covariance calculation can be simplified (similar to simplification for variance):

  $Cov(X,Y) = E(XY) - E(X)E(Y)$

- If X and Y are independent, $Cov(X,Y) = 0$.
- The converse is not true, i.e., if $Cov(X,Y) = 0$, this does *not* necessarily mean that X and Y are independent.

Covariance with the Dice Example

X: number of "2"'s rolled.

Y: sum of the numbers on the top faces.

	Y			
X	2	3	4	5
0	1/6	0	1/6	0
1	0	1/3	0	1/6
2	0	0	1/6	0

In our dice example, find $Cov(X,Y)$.

Correlation Coefficient

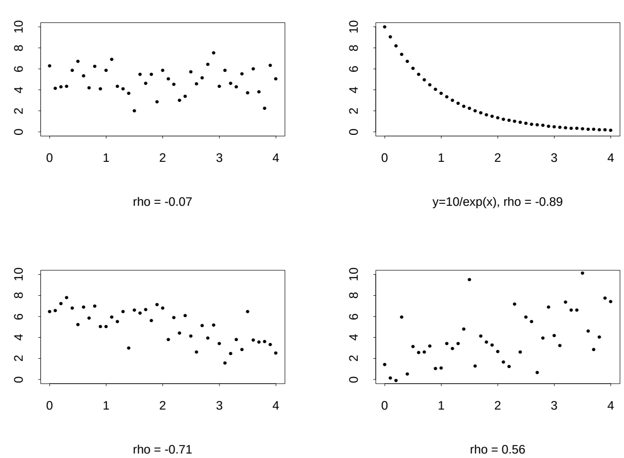
The correlation coefficient ρ between two random variables X and Y is defined (when $V(X) > 0$ and $V(Y) > 0$) as

$$\rho = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}$$

- Correlation coefficient measures strength of *linear* relationship.
- $-1 \le \rho \le 1$.
- $\rho = 1$ denotes perfect positive linear relationship (where line has positive slope).
- $\rho = -1$ denotes perfect negative linear relationship (where line has negative slope).

- $\rho = 0$ means no linear correlation, 0 covariance.
- There can be a perfect non-linear relationship between X and Y, but this won't give you $\rho = -1$ or $\rho = 1$.

How Do Various Values of ρ Look?



Correlation in the Dice Example

In the previous slide about covariance, we found

$$Cov(X,Y) = 0.25$$
$$E(X) = \frac{5}{6}$$
$$E(Y) = \frac{21}{6}$$

Find the coefficient of correlation ρ.

Variance of Linear Functions of Random Variables

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be random variables defined on the same sample space. Let $U_1 = \sum_{i=1}^{m} a_i X_i$ and $U_2 = \sum_{j=1}^{n} b_j Y_j$, where $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ are constants. Then

- $E(U_1) = \sum_{i=1}^{m} a_i E(X_i)$.
- $Cov(U_1, U_2) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j Cov(X_i, Y_j)$.
- $V(U_1) = \sum_{i=1}^{m} a_i^2 V(X_i) + \sum\sum_{i<j} a_i a_j Cov(X_i, X_j)$.

  Special Cases:

  – If X_i's are pairwise independent,then

  $V(U_1) = \sum_{i=1}^{m} a_i^2 V(X_i)$.

  – For two random variables X and Y defined on the same sample space,

  $V(aX+bY) = a^2V(X)+b^2V(Y)+2abCov(X,Y)$.