Multivariate probability distributions

- Often we are interested in more than 1 aspect of an experiment/trial
- Will have more than 1 random variable
- Interest the probability of a combination of events (results of the 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 chemical contaminants in soil samples
- Probability of obtaining a certain sample mean and sample variance in a sample from a population

Bivariate and discrete

Assume that X_1 and X_2 are discrete random variables. The joint (bivariate) probability distribution for X_1 and X_2 is $p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$.

- $p(x_1, x_2) \geq 0$ for all x_1 and x_2
- $\sum_{x_1} \sum_{x_2} p(x_1, x_2) = 1$
- C.d.f given by $F(x_1, x_2) = P(X_1 \leq x_1, Y_1 \leq x_1)$

Simple bivariate example

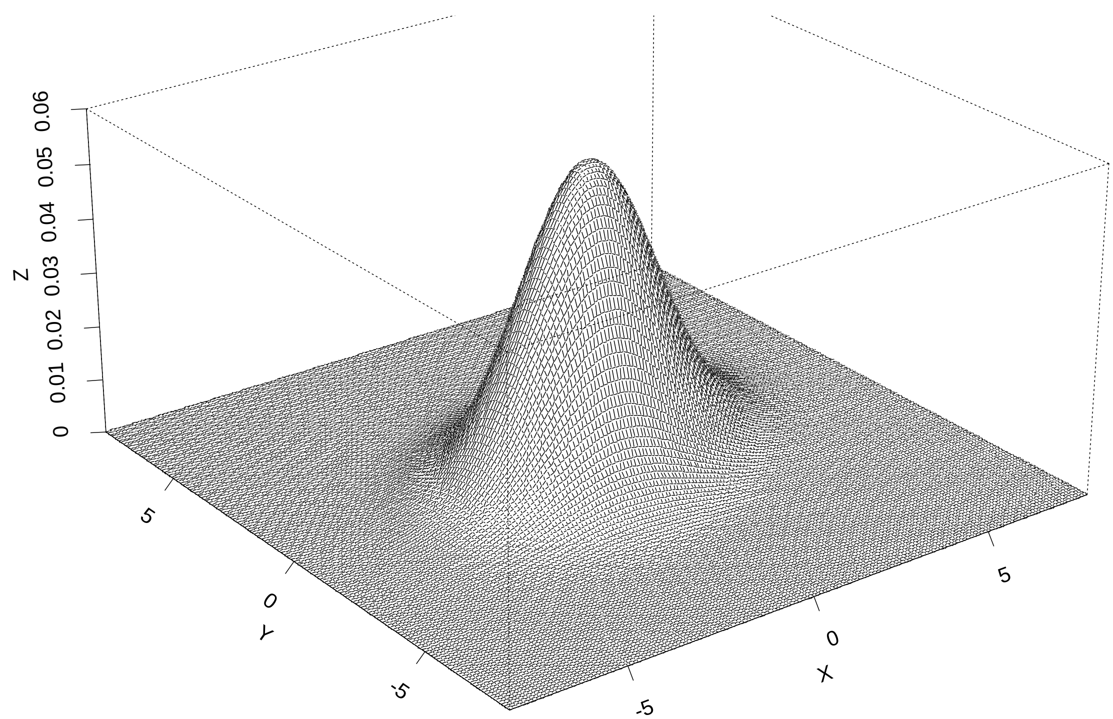
Have one die with 3 “1” faces and 3 “2” faces. Each face is equally likely to come up. The second die has 2 “1” faces, 2 “2” faces, and 2 “3” faces, also with equally weighted sides.

X: how many “2”’s rolled
Y: sum of the these dice

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

Multivariate and continuous

- Random variables Y_1 and Y_2 are jointly continuous if their joint c.d.f is continuous in both arguments
- Joint density function is given by $f(y_1, y_2)$; the function is non-negative for all y_1 and y_2
- Volume under the surface must be 1: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y_1, y_2)dy_1dy_2 = 1$
- Some calculations will require multiple integrals



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?

- Assume we have $p(x, y)$ for the random variables X and Y, want $p(x)$
- We want the probability for a union of mutually exclusive events: $(x \cap y_1) \cup (x \cap y_2) ... \cup (x \cap y_n) \rightarrow x$
- Since they're mutually exclusive, we sum the probabilities for all the different possible values of Y that can occur with $X = x$
- In the discrete case, this leads to $p(x) = \sum_y p(x, y)$ and $p(y) = \sum_x p(x, y)$
- In the continuous case, this leads to $p(x) = \int_{-\infty}^{\infty} p(x, y) dy$ and $p(y) = \int_{-\infty}^{\infty} p(x, y) dx$

Dice example continued

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

X: how many "2"'s rolled
Y: sum of the these dice

	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

Let's say we have two random variables X and Y with a joint probability function or density function, and we want to know the probability function or density function of one given the value of the other.

Discrete case:

- Use the definition of conditional probability $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- $P(X = x|Y = y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{p(x,y)}{p(y)}$

Continuous case:

- Want to have the conditional probability density function rather than a probability function
- $f(x|y) = \frac{f(x,y)}{f(y)}$

Dice example revisited

X: how many "2"'s rolled
Y: sum of the these dice

	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 revisited

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

- These 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 that A happened gives you info about $P(B)$ (or vice-versa) and the events are dependent

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

- Two discrete random variables X and Y are independent if $p(x, y) = p(x)p(y)$ for all x and y.
- Two continuous random variables X and Y are independent if $f(x, y) = f(x)f(y)$

Independence and our dice example

X: how many "2"'s rolled
Y: sum of the these dice

	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 the 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:

$$E[g(X_1, X_2, ..., X_n)] = \int_{x_1} \int_{x_2} ... \int_{x_n} g(x_1, x_2, ..., x_n) f(x_1, x_2, ..., x_n) dx_n ... dx_2 dx_1$$

Properties of expected value remain the same

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

- $E(c) = c$ where c is a constant
- $E[cg(X_1, X_2, ..., X_n)] = cE[g(X_1, X_2, ..., X_n)]$
- $E[g_1(X_1, X_2, ..., X_n) + g_2(X_1, X_2, ..., X_n)] = E[g_1(X_1, X_2, ..., X_n)] + E[g_2(X_1, X_2, ..., X_n)]$

Covariance

For two random variables X and Y:

- Denote their means as $E(X) = \mu_X$ and $E(Y) = \mu_Y$
- Covariance is an expected value:
  $Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$

The covariance calculation can be simplified (similar to simplification for variance):

Covariance

X: how many "2"'s rolled
Y: sum of the these dice

	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 $Cov(X, Y)$.

Implications of covariance

- If X and Y are independent, $Cov(X, Y) = 0$
- If $Cov(X, Y) = 0$, this does *not* necessarily mean that X and Y are independent
- It is possible for the X and Y to be dependent (according to definition, $p(x, y) \neq p(x)p(y)$), yet have $Cov(X, Y) = 0$

Coefficient of correlation

Let X be a random variable with $Var(X) = \sigma_X^2$ and Y be a random variable with $Var(Y) = \sigma_Y^2$. Then the coefficient of correlation is

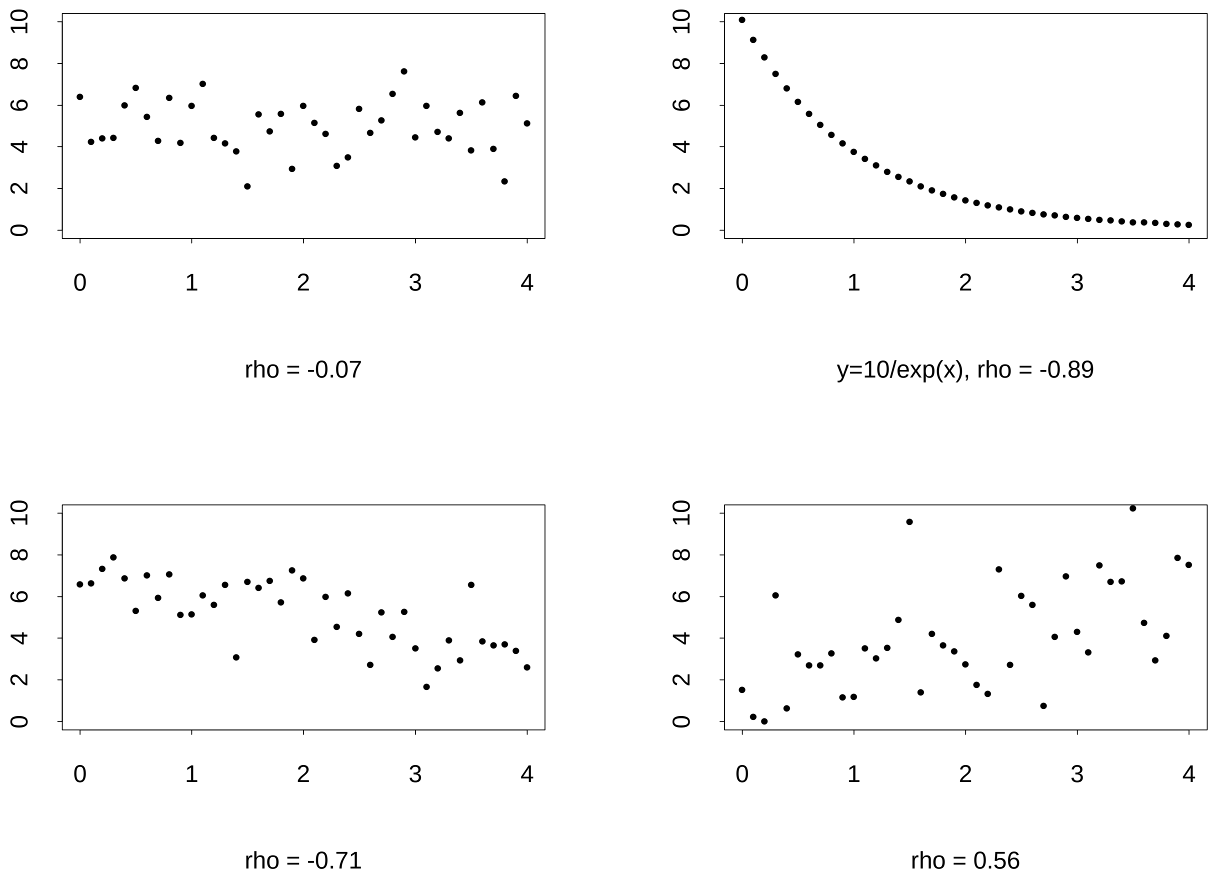
$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

- Coefficient of correlation measures strength of *linear* relationship
- $-1 \leq \rho \leq 1$, where the extremes denote perfect linear relationships (positive or negative)
- There can be a perfect non-linear relationship between X and Y, but this won't give you $\rho = -1$ or $\rho = 1$

More about ρ

- $\rho = 1$ 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 0 linear correlation, 0 covariance

How do various values of ρ look?



Correlation in the dice example

In the previous slide about covariance, we found

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

Find the coefficient of correlation ρ.

Variance for a linear function

Let's say we have two random variables X and Y, and we are interested in the variance of a linear combination $aX + bY$ of these two.

$$Var[aX + bY] = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$$

This can be extended to the case of a linear combination of $n > 2$ variables, using the same procedure. (See p. 228 for details.)