Lecture 3: Statistical Models

UNIVERSITY OF
SAN FRANCISCO

James D. Wilson

MSAN 628

Computational Statistics

Plan for this Lecture

- Random variables
- Marginal, Joint, and Conditional Probability Distributions
- Discrete vs. Continuous Random Variables
- Expectation, Variance, Covariance, and Correlation
- Independence and Mutual Independence

Random Variables

Definition

Let $(S, \mathbb{P})$ be a valid probability model. A random variable is a *real-valued function* defined on the sample space S.

Big Picture: Acts as a “measurement” of some property of a random experiment.

Notation: Use the end of the alphabet and capital letters: X, Y, Z, etc.

Examples

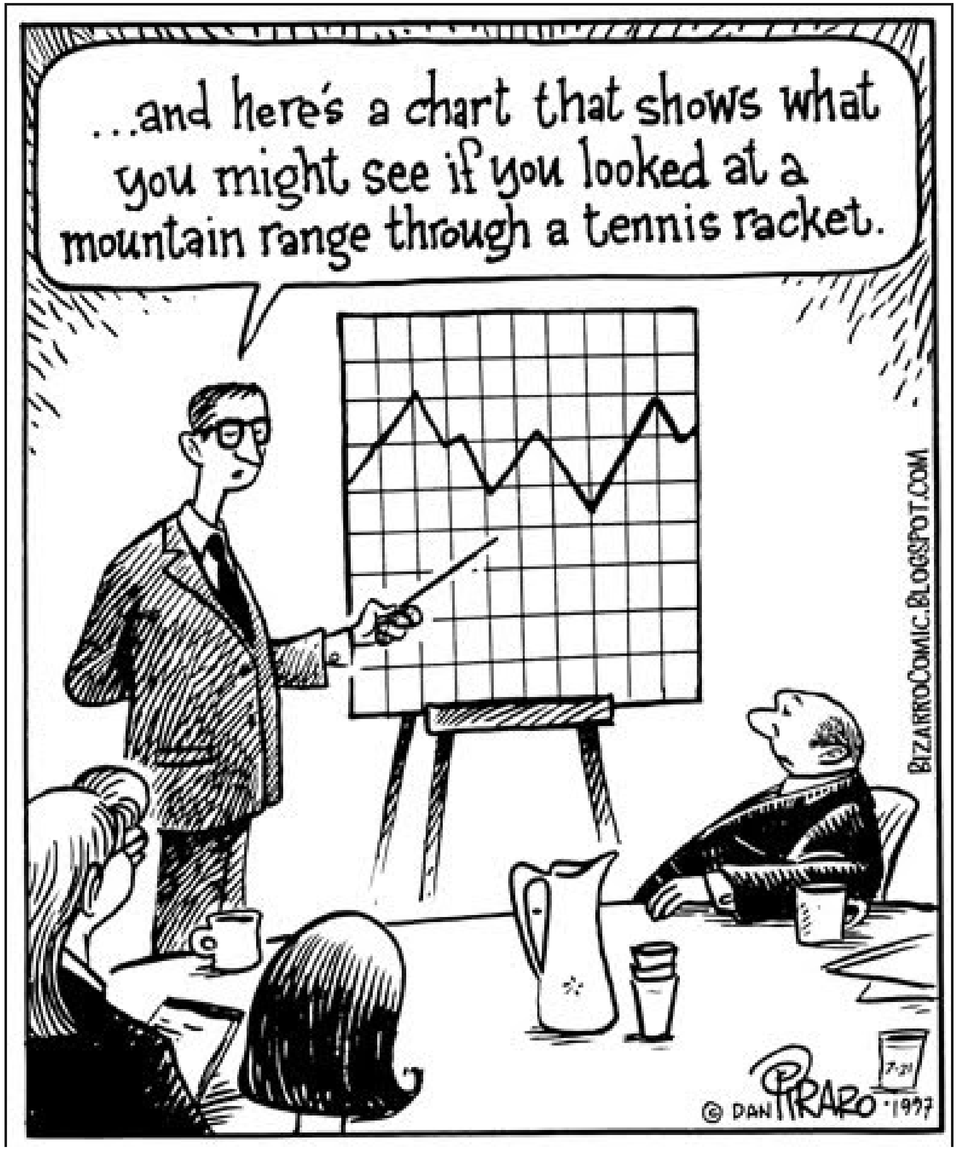
Example

Suppose that our experiment consists of tossing 3 fair coins. Let Y denote the number of heads that appear. What values can Y take?

Example

Three balls are to be randomly selected without replacement from an urn containing 20 balls numbered 1 through 20. Let X be the largest ball selected. What values can X take?

Why we need random variables...



Characterizing a Random Variable

Characterizing a random variable

Indicating what values a random variable takes and with what probabilities.

Important: there are *many* equivalent ways to characterize a random variable.

Equivalent avenues of characterization:

- probability distribution function
- cumulative distribution function
- moment generating function
- characteristic function

Any function used to characterize an RV must be one-to-one and onto.

Examples

Example

Suppose that our experiment consists of tossing 3 fair coins. Let Y denote the number of heads that appear. Characterize Y.

☼

Example

Three balls are to be randomly selected without replacement from an urn containing 20 balls numbered 1 through 20. Let X be the largest ball selected. Characterize X

☼

Types of Random Variables

Two main types of random variables Discrete and Continuous. (There are also *mixed* random variables). Treatment and analysis differ slightly.

1. For discrete, use counting arguments/summations etc.
2. For continuous, use integration and calculus etc.

Discrete random variables

X assumes one of *countably many* values: $x_1, x_2, \ldots$

$$P(X = x_i) = p(x_i) \geq 0, \quad i = 1, 2, \ldots, \quad p(x) = 0 \quad \text{for other } x.$$

Continuous random variables

X assumes values on an *uncountable* set $\mathcal{X}$.

Probability Mass Functions

Probability mass function (pmf)

Let X be a discrete random variable taking values in $\mathcal{X}$. The probability mass function of X is given by

$$p(x) = P(X = x), \qquad x \in \mathcal{X}$$

Features:

- The pmf characterizes a discrete random variable X
- A pmf $p()$ must follow the axioms of probability.
- Thus, the sum over all values must add to 1.

Random variables

Cumulative distribution function (cdf):

Let X be a random variable (continuous or discrete). The cumulative distribution function of X is:

$$F(x) = P(X \leq x), \quad -\infty < x < \infty.$$

Note: The cdf of a random variable X characterizes it as well. Why might this be?

Random variables

Example

Suppose that our experiment consists of tossing 3 fair coins. Let Y denote the number of heads that appear. What is the pmf and cdf of Y?

Continuous random variables

Continuous random variables

- The set of possible values for $\mathcal{X}$ is uncountable, such as
  - $(-\infty, \infty)$
  - $(0, \infty)$
  - (a, b), where $a, b \in \mathbb{R}$
- Examples: time until the next earthquake, the height of a randomly selected person, etc.
- Question: Now we cannot assign probabilities to each value in $\mathcal{X}$. Why not? So how can we assign probabilities?

Continuous random variables

Definition

A random variable X is a continuous random variable taking values on $\mathcal{X}$ if there is a non-negative function f on $\mathcal{X}$ such that

$$\mathbb{P}(X \in B) = \int_B f(x)dx$$

for any set $B \subseteq \mathcal{X}$.

The function f is called the probability density function (pdf) of X.

Continuous random variables

Properties

Let X be a continuous random variable taking values on $\mathcal{X} \subseteq (-\infty, \infty)$.

Then

- $\mathbb{P}(X \in \mathcal{X}) = \mathbb{P}(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f(x)dx = 1$
- $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$
- $\mathbb{P}(X = a) = \int_{\{a\}} f(x)dx = 0$

Continuous random variables

Cumulative distribution function

Let X be a continuous random variable taking values on R with probability density function f. Then the cumulative distribution function (cdf) of X is given by

$$F(a) = \mathbb{P}(X \leq a) = \mathbb{P}(X < a) = \int_{-\infty}^{a} f(x)dx, \qquad a \in \mathcal{X}$$

that is, the cdf F is the integral of the density f. Note that F is a continuous function (even if f is not).

Note

- $F'(a) = \frac{dF}{dx}(a) = f(a)$. Why is this true?

Continuous random variables

Important perspective:

For small $\epsilon > 0$,

$$\mathbb{P}\left(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}\right) = \int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x)dx \approx f(a)\,\epsilon$$

if f is continuous at $x = a$. In other words, $f(a)$ is a measure of how likely X will be near a.

Note: The above calculation also says that for a continuous random variable, for any fixed number a, the probability the random variable takes the value exactly equal to a, namely $P(X = a) = 0$.

Random Variables You Should Know

Discrete	Continuous
Poisson [Po(λ)]	Normal [N(μ, σ^2)]
Binomial [Bin(n, p)]	Uniform [U(a, b)]
Geometric [Geom(p)]	Exponential [Exp(λ)]
Hypergeometric [Hyp(N, K, n)]	Beta [Beta(a, b)]
Bernoulli [Bern(p)]	Student t (t(n))
	F [F(k, n)]
	χ^2 [$\chi^2(k)$]

Know the distribution, expectation, variance, and their applications / relationships!

Expectation

Definition

If X is a *discrete random variable* with p.m.f. $p(x)$, its expected value (or mean) is defined as

$$\mathbb{E}[X] = \sum_x xp(x).$$

Definition

If X is a *continuous random variable* with density f, its expected value (or mean) is defined as

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx.$$

Continuous random variables

Properties

- **Linearity of Expectation**: If a and b are constants, then

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

- $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ where

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx$$

- $\text{Var}(aX + b) = a^2\text{Var}(X)$
- Standard deviation: $\sqrt{\text{Var}(X)}$

Jointly distributed random variables

In this class, we will turn our attention to collections of 2 or more random variables $X_1, X_2, \ldots, X_n$. We will be interested in modeling their occurrence togther, dependence between random variables, and functions of random variables.

Examples:

- X_1 = price of stock 1, X_2 = price of stock 2, etc.
- X_1 = price today, X_2 = price yesterday, etc.
- X_1 = expenditures on food, X_2 = expenditures on housing, etc.
- X_1 = cholesterol level, X_2 = blood pressure, etc.
- X_1 = rainfall in NC, X_2 = rainfall in CA, etc.

Aim: model the joint probability of these variables together

Jointly distributed random variables

Focus

Two random variables X, Y. All probability questions about X and Y can be answered in terms of their joint cumulative distribtuion function.

The joint cumulative distribution function (joint cdf) of two random variables X and Y is given by

$$F(a,b) = \mathbb{P}(X \leq a, Y \leq b), \qquad -\infty < a, b < \infty$$

Jointly distributed random variables

Properties

- F carries info about X, Y individually:

$$F_X(a) = F(X \leq a, Y < \infty)$$

- $\mathbb{P}(X > a, Y > b) = 1 - F_X(a) - F_Y(b) + F(a,b)$
- Interval probabilities for X and Y:

$$\begin{aligned}\mathbb{P}(X \in [a_1, a_2], Y \in [b_1, b_2]) &= F(a_2, b_2) + F(a_1, b_1) \\ &\quad - F(a_1, b_2) - F(a_2, b_1)\end{aligned}$$

Jointly distributed random variables

Two broad classes of random variables:

1. **[Discrete]** Both X and Y are discrete – characterized through joint probability mass function (pmf)

$$p(x,y) = P(X = x, Y = y)$$

2. **[Continuous]** X and Y are jointly continuous: there is a non-negative function $f(x,y)$, called joint probability density function (pdf), such that, for any set C in the two-dimensional plane,

$$\mathbb{P}((X,Y) \in C) = \int\int_{(x,y)\in C} f(x,y)dxdy.$$

Joint distributions of discrete random variables

Discrete random variable

1. Characterized by their joint probability mass function

$$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$$

2. If we are given the joint pmf, then it is very easy to get the pmf of any one of the random variables. For example

$$p_X(x) = \mathbb{P}(X = x) = \mathbb{P}(X = x, \text{Y takes any value }) = \sum_y p_{X,Y}(x,y)$$

$p_X(x)$ is referred to as the marginal distribution of X.

Expectations of functions of RVs

Functions of jointly discrete random variables

If X, Y have joint pmf $p_{X,Y}$ and $g(x,y)$ is a function of the two variables (e.g $g(x,y) = x + y$ or $g(x,y) = \cos(x) + \sin(y)$) then

$$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y).$$

Functions of jointly continuous random variables

If X, Y have joint pdf $f_{X,Y}$ and $g(x,y)$ is a function of the two variables then

$$\mathbb{E}[g(X,Y)] = \int_{x,y} g(x,y) f_{X,Y}(x,y) dx dy.$$

Expectations of functions of RVs

Special case

Suppose $g(x,y) = x$. Then we get

$$\mathbb{E}[X] = \sum_{x,y} x p_{X,Y}(x,y) = \sum_{x} x \left[\sum_{y} p_{X,Y}(x,y) \right] = \sum_{x} x p_X(x).$$

Thus to calculate the expected value of X, we can

- first calculate marginal pmf p_X of X and then calculate the expected value as before $\mathbb{E}[X] = \sum_x x p_X(x)$ or
- directly calculate it using the joint pmf as above

Jointly distributed continuous random variables

Notes about jointly continuous RVs

Let X and Y be jointly continuous random variables with joint pdf $f(x,y)$. Then

- $\int\int_{(x,y)\in C} f(x,y)dxdy$ is the volume under the surface $f(x,y)$ above the region C. In particular, when $f \equiv 1$,

$$\int\int_{(x,y)\in C} dxdy = \text{Area}(C).$$

- When $C = A \times B = \{(x,y) : x \in A, y \in B\}$,

$$P(X \in A, Y \in B) = \int_A \int_B f(x,y)dydx$$

Jointly distributed continuous random variables

Notes about jointly continuous RVs

Let X and Y be jointly continuous random variables with joint pdf $f(x, y)$. Then

- For small da, db, if f is continuous at (a, b), we have

$$\mathbb{P}(a < X \leq a + da, b < Y \leq b + db)$$

$$= \int_a^{a+da} \int_b^{b+db} f(x, y) dy dx \approx f(a, b) da db$$

Thus, $f(a, b)$ is a measure of how likely (X, Y) is to be near (a, b).

Jointly distributed continuous random variables

Notes about jointly continuous RVs

Let X and Y be jointly continuous random variables with joint pdf $f(x,y)$. Then

- Each individual random variable is continuous. That is,

$$\mathbb{P}(X \in A) = \mathbb{P}(X \in A, Y \in (-\infty, \infty)) = \int_A \int_{-\infty}^{\infty} f(x,y)dydx$$

  and hence the (marginal) density of X is

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy.$$

  Similarly, $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$.

Jointly distributed random variables

More than two random variables

The notions above can be extended to more than two random variables $X_1, X_2, \ldots, X_n$. For example, the joint cdf is defined as

$$F(a_1, a_2, \ldots, a_n) = \mathbb{P}(X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n).$$

For discrete random variables we can talk about joint pmf

$$p(x_1, x_2, \ldots x_n) = \mathbb{P}(X_1 = x_1, X_2 = x_2, \cdots X_n = x_n)$$

Jointly distributed random variables

More than two random variables

In the continuous case, the random variables $X_1, X_2, \ldots, X_n$ are jointly continuous if there is a non-negative function $f(x_1, x_2, \ldots, x_n)$, called the joint probability density function (pdf), such that, for any set C in the n-dimensional space,

$$\mathbb{P}((X_1, X_2, \ldots, X_n) \in C) = \int\int_{(x_1, x_2, \ldots, x_n) \in C} \ldots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n.$$

Marginal distribution functions

Discrete case

If $(X_1, \dots X_n)$ are discrete with joint pmf p, then the marginal pmf of X_k is obtained by summing over all other dimensions, namely

$$p_{X_k}(x_k) = \sum_{x_1} \cdots \sum_{x_{k-1}} \sum_{x_{k+1}} \cdots \sum_{x_n} p(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$$

Continuous case

If $(X_1, \dots X_n)$ are continuous with joint pdf f then the marginal pdf of X_k is obtained by integrating over all other dimensions, namely

$$f_{X_k}(x_k) =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) dx_1 \cdots dx_{k-1} dx_{k+1} \cdots dx_n$$

Independence

Independence of RVs

Two random variables X and Y are independent if, for any sets A and B,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)P(Y \in B)$$

That is, the events $\{X \in A\}$ and $\{Y \in B\}$ are independent. Otherwise, we say that X and Y are dependent.

Equivalent condition 1

It can be shown that independence is equivalent to

$$F(a, b) = F_X(a)F_Y(b)$$

for all a, b

Jointly distributed random variables: Independence

Equivalent condition 2

Jointly discrete case: Independence is equivalent to

$$p(x, y) = p_X(x)p_Y(y), \quad \text{all } x, y,$$

where $p(x, y)$ is the joint pmf of X and Y, p_X is the pmf of X and p_Y is the pmf of Y.

Note: This is also equivalent to

$$p(x, y) = h(x)g(y)$$

for some functions h and g and all x, y.

Jointly distributed random variables: Independence

Equivalent condition 3

Jointly continuous case: Independence is equivalent to

$$f(x,y) = f_X(x)f_Y(y), \quad \text{all } x, y,$$

where $f(x,y)$ is the joint pdf of X and Y, f_X is the pdf of X and f_Y is the pdf of Y.

Note: This is equivalent to

$$f(x,y) = h(x)g(y), \quad \text{all } x, y,$$

for some functions h and g and all x, y.

Jointly distributed random variables: Independence

Example

If the joint density function of X and Y is

$$f(x, y) = 6e^{-2x}e^{-3y}, \quad 0 < x < \infty, \ 0 < y < \infty$$

and is equal to 0 outside this region, are the random variables independent?

What if the joint density function is

$$f(x, y) = 24xy, \quad 0 < x < 1, 0 < y < 1, 0 < x + y < 1$$

and is equal to 0 otherwise?

Independence of n random variables

Independence of n random variables

Let $X_1, \ldots, X_n$ be n random variables, either discrete or continuous. Then $X_1, \ldots, X_n$ are said to be independent if, for any sets $A_1, \ldots, A_n$,

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n)$$

$$= \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2) \ldots \mathbb{P}(X_n \in A_n)$$

In other words, the events $\{X_1 \in A_1\}, \{X_2 \in A_2\}, \ldots, \{X_n \in A_n\}$ are independent for all $A_1, \ldots, A_n$.

Sums of independent random variables:

Suppose that X and Y are two independent continuous random variables with marginal pdfs $f_X(x)$ and $f_Y(y)$ and cdfs $F_X(x)$, and $F_Y(y)$, respectively. Then the cdf of $X + Y$ is:

$$F_{X+Y}(a) = \int_{-\infty}^{\infty} F_X(a-y) f_Y(y) dy$$

Furthermore, the density of $X + Y$ is:

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$$

Name: $f_{X+Y}(a)$ is called the convolution of f_X and f_Y

Conditional distributions

Conditional distributions

Let X and Y be discrete RVs with joint pmf $p(x,y)$. Then the conditional pmf of X given $Y = y$ is:

$$p_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{p(x,y)}{p_Y(y)}$$

for all y with $p_Y(y) > 0$.

The conditional distribution function of X given $Y = y$ is

$$F_{X|Y}(a|y) = \mathbb{P}(X \leq a|Y = y) = \sum_{x \leq a} p_{X|Y}(x|y)$$

for all y with $p_Y(y) > 0$.

Conditional Distributions

Conditional distributions:

Let X and Y be jointly continuous with joint pdf $f(x,y)$. Then the conditional density of X given $Y = y$ is

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

for all y with $f_Y(y) > 0$.

The conditional probabilities of X given $Y = y$ are given by

$$\mathbb{P}(X \in A | Y = y) = \int_A f_{X|Y}(x|y)dx.$$

The conditional distribution function of X given $Y = y$ is:

$$F_{X|Y}(a|y) = \mathbb{P}(X \le a | Y = y).$$

Joint pdfs of functions of random variables

Basic problem: Suppose X_1, X_2 are jointly continuous with density $f(x,y)$.

Consider $Y_1 = g_1(X_1, X_2)$, $Y_2 = g_2(X_1, X_2)$. What is the density of Y_1, Y_2?

Transformation Theorem (Univariate)

Let X be a random variable with density $f_X(x)$. Let $Y = g(X)$ where $g(\cdot)$ is a strictly monotone function. Then,

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

Joint pdfs of functions of random variables

Transformation Theorem (Multivariate)

Let $\mathbf{X}$ be an $n-$dimensional, continuous random variable with joint density $f_{\mathbf{X}}(\mathbf{x})$ that takes values $S \subseteq \mathbb{R}^n$. Let $g = (g_1, \ldots, g_n)$ be a bijection from S to some set $T \subseteq \mathbb{R}^n$. Define

$$\mathbf{Y} = g(\mathbf{X})$$

Then the density of Y is given by

$$f_{\mathbf{Y}}(y) = \begin{cases} f_{\mathbf{X}}(g_1^{-1}(\mathbf{y}), \ldots, g_n^{-1}(\mathbf{y}))|\mathbf{J}|, & \text{for} \quad \mathbf{y} \in T \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{J}$ is the Jacobian of the tranform of X to Y, and $|\cdot|$ is the determinant.

Multivariate Properties of expectation

Property 1

Let $X_1, \ldots, X_n$ be random variables (continuous or discrete). Then

$$\mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Property 2

If X and Y are independent and $g(x), h(y)$ are two functions, then

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$$

Useful Example

Example: Expected number of events that occur

It is a common situation where we want to compute $\mathbb{E}[X]$ with X being the *number* of something. Moreover, it is often the case that for some events $A_1, A_2, \ldots, A_n$, X is the number of these events that occur (e.g. $A_i = \{\text{success on trial } i\}$). In these cases,

$$X = \sum_{i=1}^{n} I_{A_i} \quad \text{with} \quad I_{A_i} = \begin{cases} 1, & \text{if } A_i \text{ occurs,} \\ 0, & \text{if not} \end{cases}$$

and

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[I_{A_i}] = \sum_{i=1}^{n} (1 \cdot \mathbb{P}(A_i) + 0 \cdot \mathbb{P}(A_i^c)) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

Properties of Expectation

Higher-order moments of number of events that occur

For some events $A_1, A_2, \ldots, A_n$, let

$$X = \sum_{i=1}^{n} I_{A_i} \quad \text{with} \quad I_{A_i} = \begin{cases} 1, & \text{if } A_i \text{ occurs,} \\ 0, & \text{if not} \end{cases}$$

be the number of these events that occur. Note that

$$\frac{X(X-1)}{2} = \binom{X}{2} = \sum_{i_1 < i_2} I_{A_{i_1}} I_{A_{i_2}}$$

is the number of pairs of events $A_1, A_2, \ldots, A_n$ where both events occur.

Properties of expectation

Higher-order moments of number of events that occur:

More generally, the number of distinct subsets of k events that all occur can be calculated as:

$$\binom{X}{k} = \sum_{i_1 < i_2 < \ldots < i_k} I_{A_{i_1}} I_{A_{i_2}} \ldots I_{A_{i_k}}$$

is the number of distinct subsets of k events that all occur.

Hence,

$$\mathbb{E}[]\binom{X}{k}] = \sum_{i_1 < i_2 < \ldots < i_k} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k})$$

This is a *super* useful identity for "hard-to-calculate" probability problems. I've used this many times in my research.

Covariance

Now, we move onto explaining relationships *between* X and Y using expectations.

Covariance between X and Y

Let X and Y be two random variables. Then the covariance of the two random variables is

$$\text{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Idea: Covariance gives idea of the relationship between X and Y.

Covariance

Basic Properties

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, X) = \text{Var}(X)$
- $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$
- If X and Y are indepedendent, then $\text{Cov}(X, Y) = 0$

Properties of expectation

Important property

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be a collection of random variables.

Then,

$$\text{Cov}(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov}(X_i, Y_j)$$

Important consequence 1

$$\text{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \text{Var}(X_i) + 2 \sum_{i<j} \text{Cov}(X_i, X_j)$$

Important consequence 2

If X_i's are *mutually independent*, then

$$\text{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \text{Var}(X_i)$$

Correlation

Correlation between X and Y

$$\rho(X,Y) = \text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Fact 1

$$-1 \leq \rho(X,Y) \leq 1$$

Terminology: X and Y are called uncorrelated when $\rho(X,Y) = 0$.

Fact 2

Suppose that $\rho(X,Y) = -1$, then $Y = -aX + b$ with $a > 0$.

Properties of Correlation

Fact 3

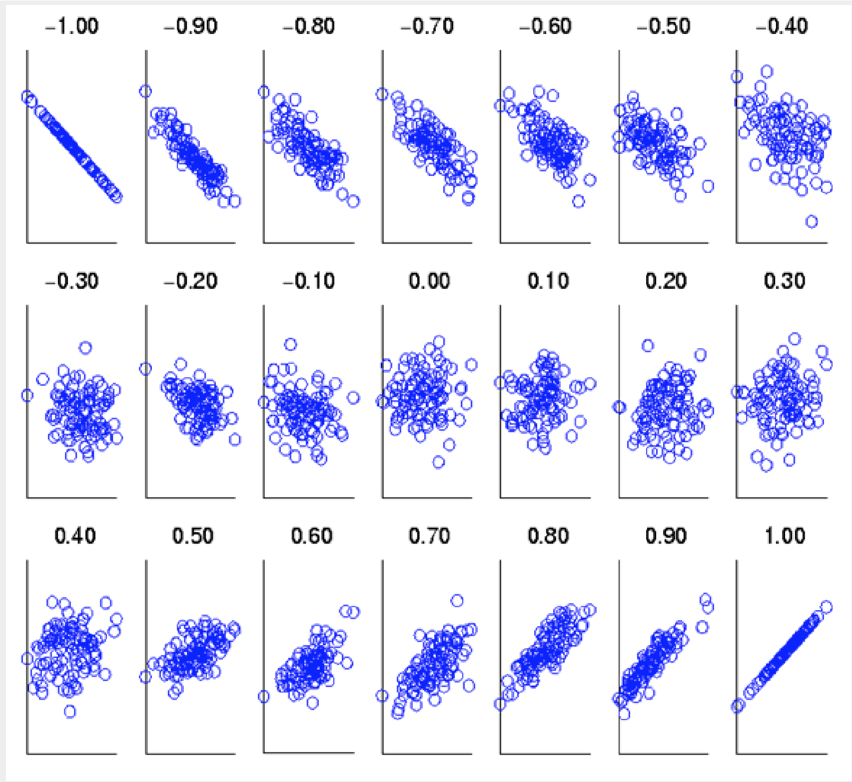
For $a > 0$,

$$\rho(aX + b, Y) = \rho(X, Y)$$

Fact 4

If X and Y are independent, then $\rho(X, Y) = 0$. But, the converse is not always true.

Properties of Correlation



Important Takeaway:

$\rho(X, Y)$ measures the strength and direction of a linear relationship between X and Y.