Probability Review

SW Chapter 2.1-2.3

EC200: Econometrics and Applications

Learning objectives

- Understand and use key vocabulary
- Calculate expected values and variances and apply their properties

Probability Review (Chapter 2.1-2.3

Key definitions: random variables

- Random variable: discrete and continuous
- Probability density function
- Cumulative density function
- Joint distribution

Random variables

Random variable — Definition

Represents a possible numerical value from a random experiment:

- Discrete random variable: Takes on no more than a countable number of values.
- Continuous random variable: Can take on any value in an interval - possible values measured on a continuum.

Discrete vs. continuous random variables

Discrete

- Roll a die twice, X is number of times 4 comes up ($X \in 0, 1, 2$).
- Toss a coin five times, X is the number of heads ($X \in 0, 1, 2, 3, 4, 5$).

Continuous

- Weight of packages filled by mechanical process
- Temperature of cleaning solution
- Time between failures of an electrical component

Probability density function

Let X be a discrete random variable and x be one of the possible values.

- The probability that X takes value x is written as $P(X = x) = P(x)$.

Probability density function — Definition

Representation of the probabilities for all possible outcomes.

- $0 \leq P(x) \leq 1$ for any value of x
- $\sum_x P(x) = 1$

Note that in the discrete case, sometimes called probability distribution function

Probability distribution function: example

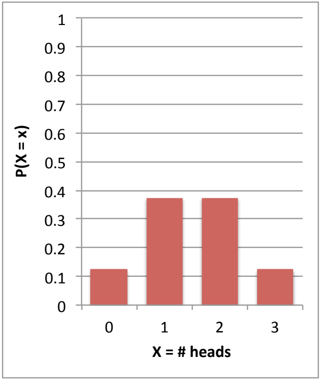
Example 1

Consider the following random experiment:

- Toss 3 coins.
- Define X as the number of heads.
- What is the probability distribution function of X? That is, show $P(x)$ for all values of x.

Probability density function: example

x	$P(x)$
0	$P(0) = 1/8 = 0.125$
1	$P(1) = 3/8 = 0.375$
2	$P(2) = 3/8 = 0.375$
3	$P(3) = 1/8 = 0.125$



Continuous random variables

- A continuous random variable has an **uncountable** number of values.
- Because there are infinite possible values, the probability of each individual value is infinitesimally small.
- If X is a continuous random variable, then $P(X = x) = 0$ for any individual value x.
- Only meaningful to talk about ranges.

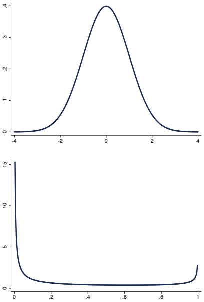
Probability density functions (PDF)

- Let X be a continuous random variable
- Its probability density function (PDF), $f(x)$ is a function that lets us compute the probability that X falls within some range of potential values.
- We define $f(x)$ such that the probability that X falls within any interval of values is equal to the *area under the curve* of $f(x)$ over that interval.

Probability density function properties

Properties of the probability density function (PDF), $f(x)$, of random variable X:

1. $f(x) > 0$ for all values of x.

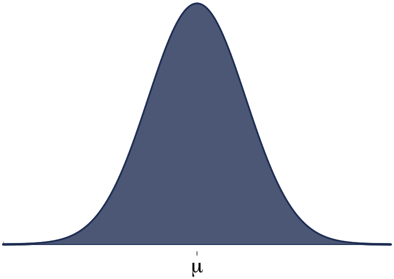


Probability density function properties

Properties of the probability density function (PDF), $f(x)$, of random variable X:

2. The area under $f(x)$ over all values of the random variable X within its range equals 1.

$$\int_X f(x)dx = 1$$

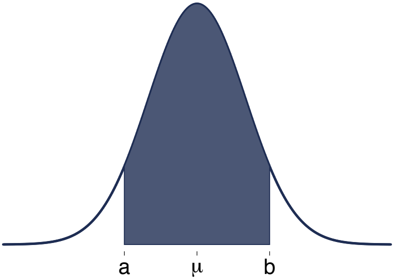


Probability density function properties

Properties of the probability density function (PDF), $f(x)$, of random variable X:

3. The probability that X lies between two values is the area under the density function graph between the two values:

$$P(a < X < b) = \int_a^b f(x)dx$$

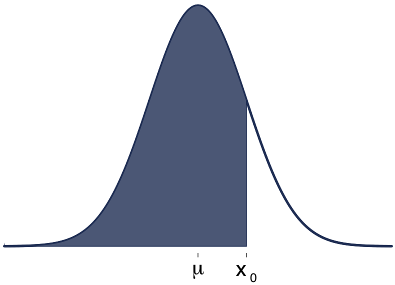


Cumulative density function (CDF)

Cumulative density function (CDF) Definition

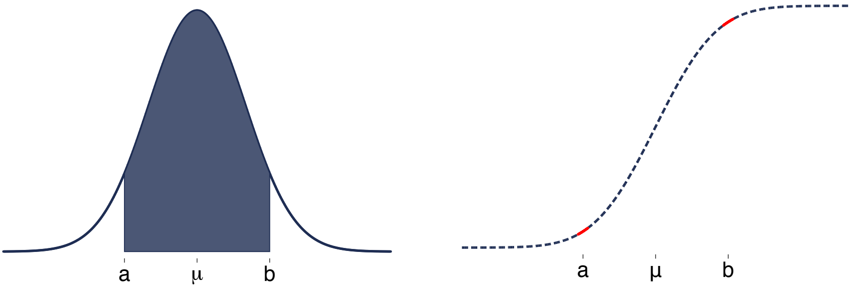
$F(x_o)$: The area under the probability density function $f(x)$ from the minimum x value up to x_0:

$$F(x_o) = \int_{x_m}^{x_0} f(x)dx$$



In some cases, $x_m = -\infty$.

Relationship between PDF & CDF



Key definitions: features of probability distributions

- Measures of central tendency: expected value
- Measures of variability: variance and standard deviation

Note: We refer to $E[Y]$ as the first moment of Y, $E[Y^2]$ as the second moment, $E[Y^3]$ as the third moment, etc.

Expected value discrete random variables

- The expected value of discrete random variable X:

$$E[X] = \mu = \sum_x xP(x)$$

- Long-run average value of the random variable X over many repeated trials
- Weighted average of possible outcomes, where weights are the probabilities of that outcome
- Also called the mean or expectation of X

Expected value of discrete random variables

Example 2

Recall an experiment in which we flip a coin 3 times. Let X be the number of heads.

X	0	1	2	3
P(x)	0.125	0.375	0.375	0.125

What is the expected value of X?

Variance/standard deviation

Variance of discrete random variable X Definition

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 P(x)$$

or

$$\sigma^2 = E[(X-\mu)^2] = \sum_x x^2 P(x) - \mu^2$$

Standard deviation of discrete random variable X Definition

$$\sigma = |\sqrt{\sigma^2}| = \sqrt{\sum_x (x-\mu)^2 P(x)|}$$

Linear functions of random variables

Let $W = a + bX$, where X has mean μ_X and variance σ_X^2, and a and b are constants:

- The mean of W is:

$$\mu_W = E[a + bX] = a + b\mu_X$$

- the variance of W is:

$$\sigma_W^2 = Var[a + bX] = b^2\sigma_X^2$$

- the standard deviation of W is:

$$\sigma_W = |b|\sigma_X$$

Joint probability distributions

What about when we have two (or more) random variables?

Joint probability distribution — Definition

Express the probability that $X = x$ and $Y = y$ simultaneously:
$P(x, y) = P(X = x \cap Y = y)$

Independence

Independence of X and Y — Definition

$$X \text{ and } Y \text{ independent } \iff P(x, y) = P(x)P(y)$$

That is, joint probability distribution is the product of their marginal probability functions for all possible values. This can be extended to k random variables

Conditional probability distributions

Conditional probability distribution Definition

The conditional probability distribution of random variable Y expresses probability that $Y = y$ conditional on $X = x$:

$$P(y|x) = \frac{P(x, y)}{P(x)}$$

Similarly,

$$P(x|y) = \frac{P(x, y)}{P(y)}$$

Conditional probability distributions: example

Example 3

The probability that the air conditioning breaks at an old factory depends on whether it is a hot day or a cold day.

- $X = 1$ if air conditioning breaks, 0 otherwise
- $Y = 1$ if it is a hot day, 0 otherwise
- Suppose $P(0,0) = 0.4$, $P(0,1) = 0.2$, $P(1,0) = 0.1$, $P(1,1) = 0.3$
- *What is the conditional marginal probability distribution of X if it is a hot day?*

Conditional probability distributions: example

	Cool day ($Y = 0$)	Hot day ($Y = 1$)
AC works ($X = 0$)	0.4	0.2
AC breaks ($X = 1$)	0.1	0.3

Conditional expectation and variance

Conditional expectation and variance Definition

We use conditional distributions to calculate the conditional expectation and conditional variance:

$$E[Y|X=x] = \sum_{i=1}^{k} y_i P(Y=y_i|X=x)$$

$$Var[Y|X=x] = \sum_{i=1}^{k} [y_i - E(Y|X=x)]^2 P(Y=y_i|X=x)$$

Covariance

- Let X and Y be discrete random variables with means μ_X and μ_Y
- The covariance between X and Y is the expected value of the product of their mean deviations

$$
\begin{aligned}
Cov(X,Y) &= E[(X-\mu_x)(Y-\mu_Y)] \\
&= \sum_x \sum_y (x-\mu_x)(y-\mu_y)P(x,y)
\end{aligned}
$$

Covariance and independence

- The covariance measures the direction of the **linear** relationship between two variables *(sometimes called "linear dependence")*.
- If two random variables X and Y are statistically independent, $\Rightarrow Cov(X, Y) = 0$.
- The converse is not necessarily true. $Cov(X, Y) = 0 \nRightarrow$ statistical independence.

Correlation

We can standardize the covariance between X and Y by dividing by their standard deviations to get the correlation between X and Y.

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

ρ is "unitless," $-1 \leq \rho \leq 1$

General rules: Linear sums and differences

Handy relationships to remember:

$$
\begin{aligned}
E[aX + bY] &= a\mu_X + b\mu_Y \\
Var(aX + bY) &= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2abCov(X, Y) \\
Var(aX - bY) &= a^2\sigma_X^2 + b^2\sigma_Y^2 - 2abCov(X, Y) \\
Cov(aX + b, cY + d) &= acCov(X, Y)
\end{aligned}
$$