Lecture 6: Discrete & Continuous Probability and Random Variables

D. Alex Hughes

Math Camp

September 17, 2015

Variance of Discrete Random Variables

Definition

The *variance* of a random variable is the expected value of its squared deviations from μ. If X is discrete with pdf $p_X(k)$,

$$VAR[X] = \sigma^2 = E[(X-\mu)^2] = \sum_{\forall k} (k-\mu)^2 \cdot p_X(k).$$

Note that $\mu = E[X]$. Also note that the standard deviation σ is the square root of the variance.

$$VAR[X] = E[(X - E[X])^2]$$

Just as the expected value has a clear analogy to the "center of balance" of a distribution, the variance is analogous to the "moment of inertia."

Imagine placing two distributions on a turntable and pushing on each with identical force 6" from the center. A distribution with low variance would spin faster, and a distribution with high variance would spin slower.

Discrete RV Variance

Example

What is the variance of a six sided die roll?

Discrete RV Variance

Example

What is the variance of a six sided die roll?

Answer

First calculate the expected value of a single die roll.

$$\sum_{1}^{6} x_i * p(x_i) = 1(1/6) + 2(1/6) + 3(1/6)$$
$$+ 4(1/6) + 5(1/6) + 6(1/6)$$
$$E[X] = \mu = 3.5$$

Discrete RV Variance

Answer

Then the Variance is:

$$\sum_{1}^{6}(x_i - \mu)^2 p(x_i) = (1-3.5)^2(1/6) + (2-3.5)^2(1/6) + (3-3.5)^2(1/6)$$
$$+ (4-3.5)^2(1/6) + (5-3.5)^2(1/6) + (6-3.5)^2(1/6)$$
$$= 17.5(1/6)$$
$$= 2.92$$

Example of Discrete RV Variance

Example

An urn contains five chips, two red and three white. Two chips are drawn at random, *with replacement*. Let X denote the number of red chips drawn. What is VAR[X]?

Example of Discrete RV Variance

Example

An urn contains five chips, two red and three white. Two chips are drawn at random, *with replacement*. Let X denote the number of red chips drawn. What is VAR[X]?

Answer

X is binomial, $n = 2$ and $p = \frac{2}{5}$.

$$
\begin{aligned}
E[X] &= np = (2)(2/5) = 4/5 \\
VAR[X] &= \sum (x - E[X])^2 f(x) = \sum_{i=0}^{2} (i - .8)^2 \binom{2}{i} \left(\frac{2}{5}\right)^i \left(\frac{3}{5}\right)^{2-i} \\
&= (0 - .8)^2 \cdot \binom{2}{0} (2/5)^0 (3/5)^2 + (1 - .8)^2 \cdot \binom{2}{1} (2/5)^1 (3/5)^1 + (2 - .8)^2 \binom{2}{2} (2/5)^2 (3/5)^0 \\
&= (16/25)(1)(9/25) + (1/25)(2)(6/25) + (36/25)(1)(4/25) = \mathbf{\frac{12}{25}}
\end{aligned}
$$

Example

Example

Some process has equal chance of generating a number between [0,1]. What is the probability the process generate a number less that 0.5?

Example

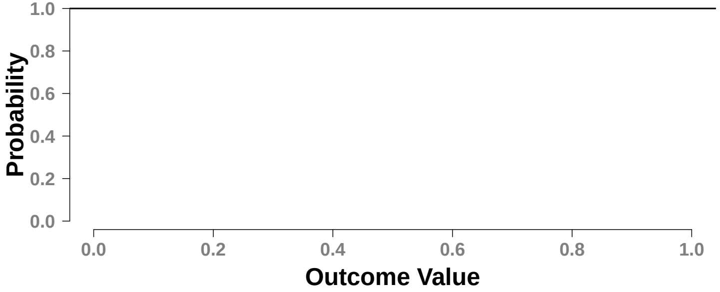
Example

Some process has equal chance of generating a number between [0,1]. What is the probability the process generate a number less that 0.5? Draw a number line.

Example

Example

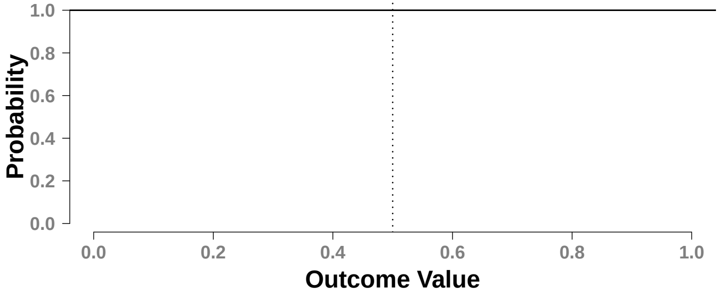
Some process has equal chance of generating a number between [0,1]. What is the probability the process generate a number less that 0.5? Draw a number line.



Example

Example

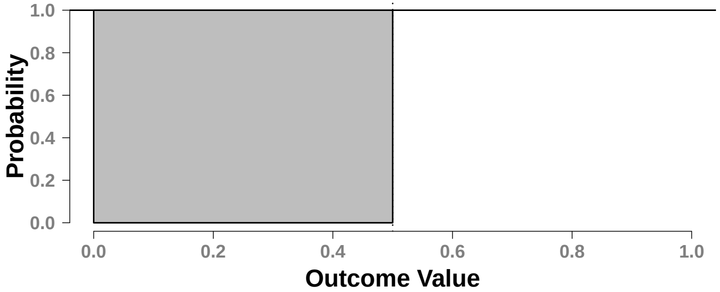
Some process has equal chance of generating a number between [0,1]. What is the probability the process generate a number less that 0.5? Draw a number line.



Example

Example

Some process has equal chance of generating a number between [0,1]. What is the probability the process generate a number less that 0.5? Draw a number line.



Continuous Probability

Definition

A probability density function P on a set of real numbers S is called *Continuous* is there exists a function f(t) such that for any closed interval [a,b] $\subset$ S,

$$P([a,b]) = \int_a^b f(t)dt$$

Continuous Probability

Definition

A probability density function P on a set of real numbers S is called *Continuous* is there exists a function f(t) such that for any closed interval [a,b] $\subset$ S,

$$P([a, b]) = \int_a^b f(t)dt$$

Comment

f(t) must satisfy two properties:

1. $f(t) \geq 0$ for all t
2. $\int_{-\infty}^{\infty} f(t)dt = 1$

Frequently, we refer to the probability distribution function as the "pdf."

Continuous Probability Example

Example

- What continuous probability function describes the uniform chance of drawing a number on the range [0,10]?
  - What is the probability that a randomly selected number falls between 4 and 7?
  - What is the probability that a randomly selected number falls between [2,4] or [6,9]?
- What continuous probability function describes the uniform chance of drawing a number on the range [a,b]?

Continuous Probability Example 2

Example

Is $f(x) = 3x^2, 0 \leq x \leq 1$ a valid probability function?

Continuous Probability Example 2

Example

Is $f(x) = 3x^2, 0 \leq x \leq 1$ a valid probability function?

- What is the probability that $0 \leq x \leq 1/3$?
- What is the probability that $2/3 \leq x \leq 1$?

Continuous Probability Example 2

Example

Is $f(x) = 3x^2, 0 \leq x \leq 1$ a valid probability function?

- What is the probability that $0 \leq x \leq 1/3$?
- What is the probability that $2/3 \leq x \leq 1$?

Answer

- Yes. Recall the two properties f(x) just satisfy: (A) f(x) is always non-negative; (B) The integral of f(x) across the function's domain = 1.
- $\int_0^{1/3} = x^3|_0^{1/3} = \mathbf{1/27}$
- $\int_{2/3}^{1} = x^3|_{2/3}^{1} = 1 - 8/27 = \mathbf{19/27}$.

Example 3

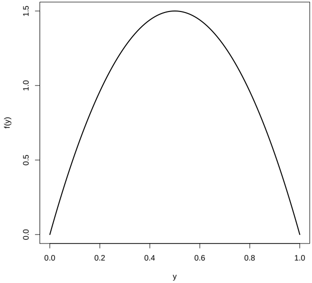
Example

Suppose we would like a continuous function to describe a variable Y in such a way that it more likely to produce y's near the middle of the range more frequently than near the extremes of the range. What might this look like?

Example 3

Example

Suppose we would like a continuous function to describe a variable Y in such a way that it more likely to produce y's near the middle of the range more frequently than near the extremes of the range. What might this look like?



I've drawn one where $f(y) = 6y(1 - y)$.

Example 3

Example

Does this meet the requirements of a probability density function?

Example 3

Example

Does this meet the requirements of a probability density function?

Answer

Yes!

1. $f_Y \geq 0$ for all y;
2. $P(\Omega) = 1$.

Then, with this function, answer the following:

- What is the probability that a number is less than 0.5?
- Greater than 0.5?
- Less than 0.25?
- Greater than 0.25?
- Exactly 0.25?

Cumulative Density Function

Definition

The cumulative density function (called "cdf") of a continuous, real-valued random variable Y is an indefinite integral of its pdf.

$$F_Y(x) = \int_{-\infty}^{x} f_y(x)dx = P(X \leq x)$$

Note the difference in notation between a cumulative distribution function (cdf) and a probability distribution function (pdf).

Theorem

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(t)dt; \textit{ conversely,} \\ \frac{d}{dy}F(y) &= f(y) \end{aligned}$$

Cumulative Density Examples

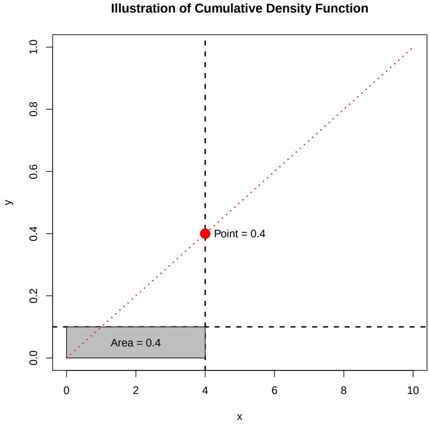
Example

A uniform distribution covers the range [0,10]. What is the probability that a random number drawn is less than 4?

Cumulative Density Examples

Example

A uniform distribution covers the range [0,10]. What is the probability that a random number drawn is less than 4?



$$\int_0^4 f(x)dx = .1x\Big|_0^4$$
$$= .1(4) - .1(0)$$
$$\int f(x) = F_X(x)$$
$$F_X(x) = 0.1x$$
$$F_X(4) = 0.1(4)$$
$$= .4$$

CDF Properties

Theorem

Let Y be a continuous random variable with cdf $F_Y(y)$, Then,

1. $P(Y > s) = 1 - F_Y(s)$
2. $P(r < Y \leq s) = F_Y(s) - F_Y(r)$
3. $\lim_{x \to +\infty} F_Y(y) = 1$
4. $\lim_{x \to -\infty} F_Y(y) = 0$

Additionally:

- The CDF is a monotonically nondecreasing continuous function.
- Will be useful for easily/quickly calculating the probability a variable takes a value on some interval.

An Example

In Encinitas the distribution of waves a broski, Tad, takes every hour, Y, is described by the pdf:

$$f_Y(y) = y * e^{-y}, y \geq 0.$$

What is the median of Encinitas' wave-income distribution?

An Example

In Encinitas the distribution of waves a broski, Tad, takes every hour, Y, is described by the pdf:

$$f_Y(y) = y * e^{-y}, y \geq 0.$$

What is the median of Encinitas' wave-income distribution?

In looking for the median, we are looking for the number where the probability below is equal to the probability above.

An Example

In Encinitas the distribution of waves a broski, Tad, takes every hour, Y, is described by the pdf:

$$f_Y(y) = y * e^{-y}, y \geq 0.$$

What is the median of Encinitas' wave-income distribution?

In looking for the median, we are looking for the number where the probability below is equal to the probability above.

$$F_Y(y) = \int_0^\infty ye^{-y}$$
$$= -e^{-y}$$
$$0.5 = -e^{-y}|_y^\infty$$
$$0.5 = 0 - (-e^{-y})$$
$$0.5 = e^{-y}$$
$$ln(0.5) = -y \to \mathbf{0.69}$$

Definition

If Y is a continuous random variable with pdf $f_Y(y)$,

$$E[Y] = \int_{-\infty}^{\infty} y \cdot f_Y(y)dy$$

In the same was as the discrete case, one may think of this as a center of gravity of the distribution.

Example

A continuous uniform density function, f(x) = 1 is defined on the range [0,1]. What is the expected value of this function?

Answer

$$
\begin{aligned}
E[X] &= \int_0^1 x \cdot f(x) dx \\
&= \int_0^1 1x dx \\
&= \frac{1}{2} x^2 |_0^1 \\
&= \frac{1}{2}(1 - 0) \\
&= \mathbf{0.5}
\end{aligned}
$$

Variance of Continuous Random Variables

Definition

Let X be a real-valued random variable with density function $f_X(x)$. Then, the *variance*, σ^2, is defined by

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

Expectation Identities

Theorem

For any constant c, $E[c] = c$

Expectation Identities

Theorem

For any constant c, $E[c] = c$

Theorem

For any random variable W, $E[W + b] = E[W] + b$

Theorem

For any constant c, $E[c] = c$

Theorem

For any random variable W, $E[W + b] = E[W] + b$

Theorem

For any constant c and random variable X,

$$E[cX] = cE[X]$$

Expectation Identities

Theorem

For any constant c, $E[c] = c$

Theorem

For any random variable W, $E[W + b] = E[W] + b$

Theorem

For any constant c and random variable X,

$$E[cX] = cE[X]$$

Theorem

For any X and Y that are random variables that have finite expectations

$$E[X + Y] = E[X] + E[Y]$$

Expected Value of a Function of a Random Variable

Theorem

Suppose X is a random variable with pdf $p_X(k)$. Let $g(X)$ be any function of X. Then the expected value of the random variable g(X) is given by

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx$$

Example

Suppose $X = f_X(x)$ is the uniform density function defined on the range [0,1]. Then $f_X(x) = 1$. Suppose $g(x) = 1/5x$.

$$E[g(X)] = \int_0^1 \frac{1}{5} x \cdot f_X(x) dx$$
$$= \frac{1}{5} [\frac{1}{2} x^2]_0^1 = \mathbf{\frac{1}{10}}$$

Example

The Zetas want to burn their gang-signs into a wall. Suppose the amount of fuel in the torch, Y, is a random variable with pdf

$$f_Y(y) = 3y^2, 0 < y < 1.$$

In the past they have been able to burn a circle whose radius is 20 times the size of y. How much area can they expect to burn?

Example

The Zetas want to burn their gang-signs into a wall. Suppose the amount of fuel in the torch, Y, is a random variable with pdf

$$f_Y(y) = 3y^2, 0 < y < 1.$$

In the past they have been able to burn a circle whose radius is 20 times the size of y. How much area can they expect to burn?

Answer

By problem setup, $g(Y) = 20\pi Y^2$

$$E[g(Y)] = \int_0^1 20\pi y^2 \cdot 3y^2 dy = 60\pi \int_0^1 y^4$$
$$= \frac{60\pi y^5}{5}|_0^1$$
$$= 12\pi ft^2$$

Variance

Recall the definition of Variance:

Definition

The *variance* of a random variable is the expected value of its squared deviations from μ. If X is discrete with pdf $p_X(k)$,

$$VAR[X] = \sigma^2 = E[(X-\mu)^2] = \sum_{\forall k}(k-\mu)^2 \cdot p_X(k).$$

Note that $\mu = E[X]$. Also note that the standard deviation σ is the square root of the variance.

$$VAR[X] = E[(X - E[X])^2]$$

A Variance Identity

Theorem

Let X be any random variable (discrete or continuous), having mean μ and for which $E[X^2]$ is finite. Then,

$$VAR[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

Proof.

$$\begin{aligned} VAR[X] &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2] \\ &= E[X^2] - E[2XE[X]] + E[E[X]^2] \\ &= E[X^2] - E[2X]E[X] + E[X]^2 \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

An Example

Example

Given the following pdf, what is the variance?

$$f(x) = 3(1-x)^2, 0 < x < 1$$

An Example

Example

Given the following pdf, what is the variance?

$$f(x) = 3(1-x)^2, 0 < x < 1$$

Answer

$$E[X] = \int_0^1 x \cdot 3(1-x)^2 dx = 3\int_0^1 x - 2y^2 + y^3 = 3$$

$$E[X^2] = \int_0^1 x^2 \cdot 3(1-x)^2 dx = 3\int_0^1 y^2 - 2y^3 + y^4 dx = \frac{1}{10}$$

$$VAR[X] = E[X^2] - E[X]^2 = \frac{3}{80}$$

Variance Identities

Theorem

For and random variable X and constant c

$$VAR[cX] = c^2 VAR[X]$$

$$VAR[X + c] = VAR[X]$$

Proof.

For ease, let $E[X] = \mu$. Then $E[cX] = c\mu$, and

$$\begin{aligned} V[cX] &= E[(cX - c\mu)^2] = E[c^2(X - \mu)^2] \\ &= c^2 E[(X - \mu)^2] \\ &= c^2 VAR[X] \end{aligned}$$

□

Variance Identities

Theorem

Let X and Y be two independent random variables. Then,

$$VAR[X+Y] = VAR[X] + VAR[Y]$$

Proof.

Let E[X] = a and E[Y] = b.

$$\begin{aligned} VAR[X+Y] &= E[(X+Y)^2] - (a+b)^2 \\ &= E[X^2] + 2E[XY] + E[Y^2] - a^2 - 2ab - b^2 \\ &= E[X^2] - a^2 + E[Y^2] - b^2 = VAR[X] + VAR[Y] \end{aligned}$$

□

Introduction

What do we require of a probability distribution/density model?

- It must be non-negative for all outcomes in its sample space
- It must sum or integrate to one across the sample space.

By this definition, $f(y) = \frac{y}{4} + 7\frac{y^3}{2}, 0 \leq y \leq 1$ is a valid pdf.

- $f(y) \geq 0, \forall y 0 \leq 1$, and;
- $\int_0^1 \frac{y}{4} + 7\frac{y^3}{2} dy = 1$

But, how useful is this model?

A Model's Merit

- A pdf has utility as a probability model if it actually models the behavior of the real (or political) world.
- A surprisingly small number of distributions describe these real world outcomes.
- Many measurements/outcomes are the result of the same set of assumptions about the data generating process
    - Number of fraud incidents
    - Number of latrines built
    - Feeding patterns of zebra muscles