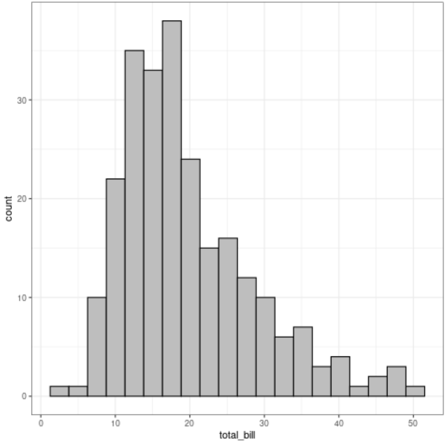
Probabilitity: Random Variables

Probabilitity Distributions, Expectation, and Variance

Grinnell College

Review – Describing Distributions

When we saw *distributions* earlier in the class, they were a way to represent information from a quantitative variable.



Two of the big things we talked about were center and spread.

- Measures of center we used were median or mean
- Measures of spread we used were IQR or std. dev.

Review – Describing Distributions

Mean is the same thing as the **average** value of the variable.

$$\overline{x} = \frac{\sum x_i}{n}$$

- even if a distribution is skewed, the mean can still be useful
  - more on this in a sec

Standard Deviation is one of the measures of spread

$$s = \sqrt{\frac{1}{n-1} \sum (x_i - \overline{x})^2}$$

- interpretation: the average distance of observations from the mean
- larger value of s $\rightarrow$ more variability or spread
- sometimes variance is used instead (variance $= s^2$)

Goal for Today

We are going to apply the concept of center and spread (mean and standard deviation) to the probability distributions concept we saw previously.

- give names to certain shapes
- some recurring probability situations

Review – Probability Distributions

A *discrete* **probability distribution** represents each of the *disjoint* outcomes of a random process and their associated probabilities

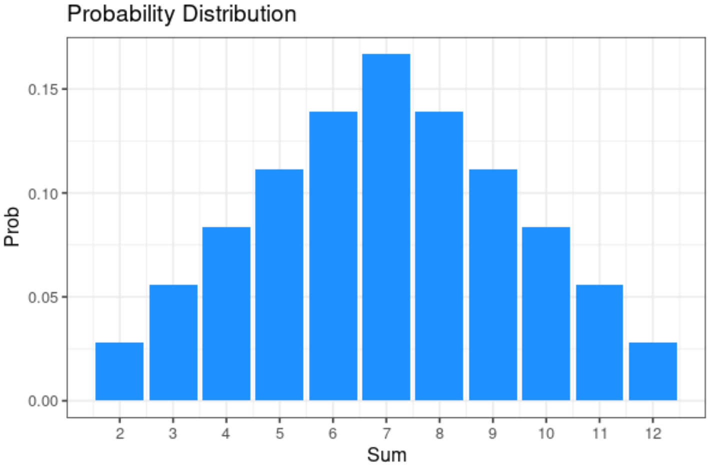
Dice Sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

For a *discrete* probability distribution to be valid, the following must be true:

1. The outcomes are disjoint
2. Every probability is between 0 and 1
3. The sum of all probabilities must equal 1

Review – Probability Distributions

Dice Sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$



Random Variable

When working with a *random process* (like die rolling, or coin flipping) we can construct a quantitative variable that tells us about the outcome of that process.

Typically we will label a random variable with capital letters to distinguish it from data variables in our data sets.

Example 1: rolling a six-sided die

- X = result of die roll (can be 1,2,3,4,5 or 6)
- $P(X=6) = \frac{1}{6}$

Example 2: coin flip

- Y = 1 if heads, Y = 0 if tails
- $P(Y=1) = P(H) = P(T) = P(Y = 0) = 0.5$

Expectation

When talking about the center of a probability distribution, most often the *mean* is used (even if the distribution is skewed). We will use the term **Expected Value** to denote that this is an average for a *random variable*.

To compute the **Expected Value**, we need to use the outcomes and account for how likely they are to come up. The *expected value* of a random variable X is denoted E(X)

Formula:

$$E(X) = \sum x_i p_i$$

- x_i represent the value of outcome i
- p_i is the probability associated with outcome i

Expectation – Example

Let's look at the probability distribution for rolling one 6-sided die.

Die Result	1	2	3	4	5	6
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Let X = result of roll, then

$$\begin{aligned} E(X) &= \sum x_i p_i \\ &= 1(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + 4(\frac{1}{6}) + 5(\frac{1}{6}) + 6(\frac{1}{6}) \\ &= \frac{21}{6} = 3.5 \end{aligned}$$

- Interpretation: the *expected* outcome is 3.5
- Interpretation: If you roll many 6-sided dice and compute the average, you can expect a value close to 3.5

Variance

We may also want to talk about the *variability* of a process. It's nice to know the average of a d6 is 3.5, but how much variability can I expect when I roll it?

Working with **variance** ($=s^2$) is usually easier than std.dev. directly, although interpretations with std.dev. are easier

- variance of a random variable is denoted Var(X)

Formula: Let $\mu = E(X)$.

$$Var(X) = E[(X-\mu)^2] = E(X^2) - \mu^2$$
$$= (\sum x_i^2 p_i) - \mu^2$$

- variance = expected squared deviation from the mean
- we won't do much calculation of Variance directly
- convert to std.dev. to do interpretations

Variance Example

Let's go back to the 6-sided die rolling example.

Die Result	1	2	3	4	5	6
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Let X = result of roll, then

$$\begin{aligned} Var(X) &= (\sum x_i^2 p_i) - \mu^2 \\ &= [1^2(\frac{1}{6}) + 2^2(\frac{1}{6}) + 3^2(\frac{1}{6}) + 4^2(\frac{1}{6}) + 5^2(\frac{1}{6}) + 6^2(\frac{1}{6})] - 3.5^2 \\ &= \frac{91}{6} - 3.5^2 = \frac{105}{36} \approx 2.92 \end{aligned}$$

- variance = 2.92 $\rightarrow$ s.d. = $\sqrt{2.92} = 1.71$
- interpretation: a d6 result is 1.71 away from the mean, on average

Helpful: Expectation for Multiple RV's

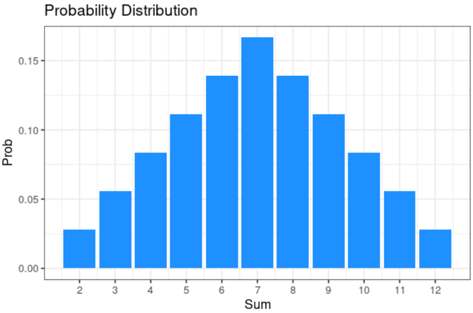
Let's say I wanted to calculate the expected value when I'm working with multiple dice. When computing the expected value of the sum of things, it can be time consuming to keep track of all the outcomes. There is a simpler rule.

Expectation of Sums

Let X and Y be two random variables. Their sum $X + Y$ has expected value $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$

This works for more than 2 RV's too. The rule makes finding an average of lots of processes very easy

Sum of Dice Example



For rolling 2 dice and adding them up, visually we can see the expected value (mean) should be 7. Let's use the result from the last slide

- Let X = result of die 1
- Let Y = result of die 2
- Let Z = sum of two dice = X + Y
- Then $E(Z) = E(X) + E(Y) = 3.5 + 3.5 = 7$

Helpful: Variance Rules

There are also rules for variances similar to expected values when working with multiple random variables.

- Var(X $\pm$ Y) $=$ Var(X) $+$ Var(Y) if X and Y are independent
- Var(X - Y) $=$ Var(X) $+$ Var(Y) if X and Y are independent
- Var(cX) $=$ c^2Var(X)

When X and Y associated:
Var(aX $\pm$ bY) $=$ a^2Var(X) $+$ b^2Var(Y) $\pm$ 2abE[(X-μ_X)(Y-μ_Y)]

More on Random Variables

Definition: A *Random Variable* is a variable in which the value is determined by a random event. Often these are labeled with capital letter like X and Y.

If we want to be completely thorough, we need to define the sample space $\mathbf{\Omega}$ (set of all possible outcomes) and the probabilities for each.

There are certain situations which show up repeatedly in probability applications, and so we will give these distributions special names. These are governed by *parameters*, numbers which define a probability structure.

Parameter: A number which determines part or all of a probability distribution.

Bernoulli Distribution

X $\sim$ Bernoulli(p) (parameter = p)

Use: Modeling binary events. Probability of success is defined as p. Let the outcome "1" denote success, "0" failure.

$$\mathrm{P(X{=}1) = p} \qquad \mathrm{P(X{=}0)} = 1 - p := \mathrm{q}$$

$$P(X = k) = \begin{cases} p & k = 1 \\ q = 1 - p & k = 0 \end{cases}$$

$\mathrm{E(X)} = \sum xP(X = x)$
$= \sum_{x=0}^{1} xP(X = x) = \mathrm{p}$

$\mathrm{Var(X) = E(X^2)} - E(X)^2$
$= \sum x^2 P(X = x)$ - $\mathrm{p}^2 = pq$

Example: Fair Coin E(X) = 0.5, Var(X) = 0.25

Binomial Distribution

X $\sim$ Binomial(n, p)
Use: Models the number of successes in n independent Bernoulli trials, each with probability of success p.

$$\mathrm{P(X = k)} = \binom{n}{k} p^k (1-p)^{n-k} \qquad \mathrm{k = 0, 1, \ldots, n}$$

$$\mathrm{E(X)} = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np \qquad \mathrm{Var(X)} = np(1-p)$$

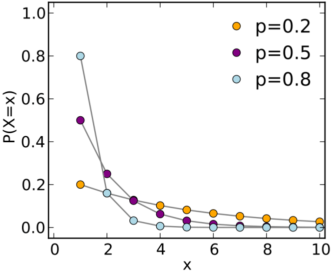
Example: Rolling a fair die 10 times, counting 6's.
$\mathrm{E(X)} = 10 \times \frac{1}{6} = 1.67, \quad \mathrm{Var(X)} = 10 \times \frac{1}{6} \times \frac{5}{6} \approx 1.39$

Geometric Distribution

X $\sim$ Geometric(p)

Use: Models the number of trials until the first success in independent Bernoulli trials, each with success probability p.

$$P(X = k) = (1 - p)^{k-1} p \qquad k = 1, 2, 3, \ldots$$



$$E(X) = \sum_{k=1}^{\infty} k(1-p)^{k-1} p = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$

Tossing a biased coin ($p = 0.6$), counting flips until first heads:

$$E(X) = \frac{1}{0.6} \approx 1.67, \qquad Var(X) = \frac{0.4}{0.6^2} \approx 1.11$$

Discrete Uniform Distribution

X $\sim$ unif{a,b}
Use: Models equally likely outcomes over the integers from a to b.

$$\mathrm{P(X = k)} = \begin{cases} \dfrac{1}{b-a+1}, & k = a, a+1, \ldots, b \\ 0, & \text{otherwise} \end{cases}$$

$$\mathrm{E(X)} = \frac{a+b}{2} \qquad \mathrm{Var(X)} = \frac{(b-a+1)^2-1}{12}$$

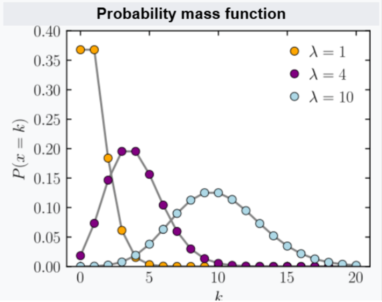
Example: Roll of a fair die: $X \sim \mathrm{unif}\{1,6\}$
E(X) = 3.5, Var(X) = 2.92

Poisson Distribution

X $\sim$ Poisson(λ)

Use: Models the number of events occurring in a fixed interval of time, assuming events occur independently at an average rate λ.

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad k = 0, 1, 2, \ldots$$



$E(X) =$

$\sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} = \lambda$

$Var(X) =$

$\sum_{k=0}^{\infty} (k^2 - \lambda^2) \frac{e^{-\lambda}\lambda^k}{k!} = \lambda$

Example: Number of cars passing a checkpoint in one minute, if $\lambda = 4$:
$E(X) = 4, \quad Var(X) = 4$

Continuous Random Variables

A *continuous* **random variable** can take on values on the real line (not just discrete points). Its probabilities are described by a **probability density function** (pdf) $f(x)$ rather than by a list of discrete probabilities.

For a *continuous* probability distribution to be valid:

1. $f(x) \geq 0$ for all x
2. The total probability is 1: $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Expectation:

$$E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx$$

Variance:

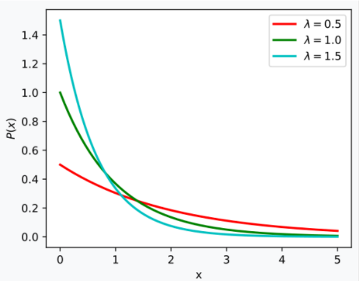
$$Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

where $\mu = E(X)$

Exponential Distribution

X $\sim$ Exponential(λ)
Use: Models the time between independent events occurring at a constant average rate λ.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



Expectation:

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}\,dx = \frac{1}{\lambda}$$

Variance:

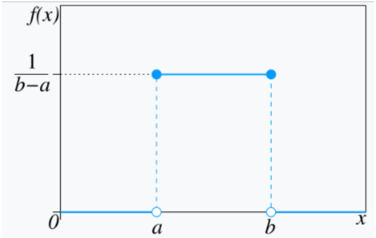
$$Var(X) = \int_0^\infty (x-\mu)^2 \lambda e^{-\lambda x}\,dx = \frac{1}{\lambda^2}$$

Continuous Uniform Distribution

X $\sim$ Uniform(a, b)

Use: Outcomes that are equally likely over the continuous interval $[a, b]$.

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$



Expectation:

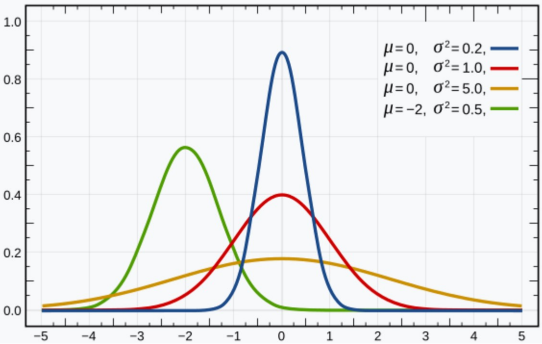
$$E(X) = \int_a^b x \frac{1}{b-a}\, dx = \frac{a+b}{2}$$

Variance:

$$Var(X) = \int_a^b (x-\mu)^2 \frac{1}{b-a}\, dx = \frac{(b-a)^2}{12}$$

Normal Distribution

$X \sim \text{Normal}(\mu, \sigma^2)$



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

Expectation:

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

Variance:

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