

Bernoulli random variables

Earlier we defined a binomial random variable as a variable that takes on the discrete values of “success” or “failure.” For example, if we want heads when we flip a coin, we could define heads as a success and tails as a failure. We could model this scenario with a binomial random variable X where X is the number of times we get heads when we flip a coin a specified number of times.

A **Bernoulli random variable** is a special category of binomial random variables. Specifically, with a Bernoulli random variable, we have exactly one trial only (binomial random variables can have multiple trials), and we define “success” as a 1 and “failure” as a 0.

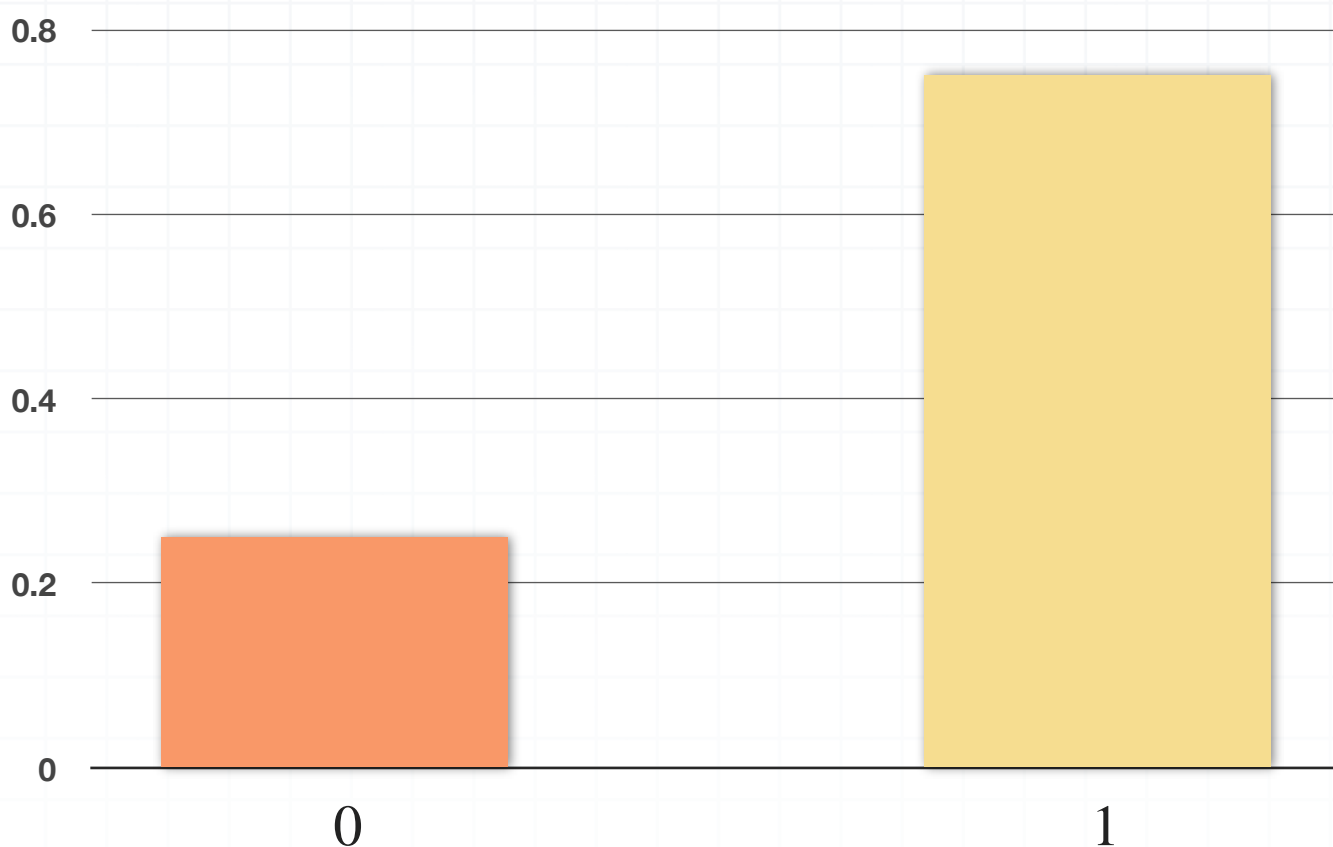
Bernoulli distributions

Let’s say I want to know how many students in my school like peanut butter. I can’t survey the entire school, so I survey only the students in my class, using them as a sample. I ask them whether or not they like peanut butter, and I define “liking peanut butter” as a success with a value of 1 and “disliking peanut butter” as a failure with a value of 0. I find that 75 % of the students in my class like peanut butter.

Since everyone in our survey was forced to pick one choice or the other, 100 % of our population is represented in these two categories, which means that the probability of both options will always sum to 1.0 or 100 %. Therefore, since 75 % of the students in my class like peanut butter, that means $100\% - 75\% = 25\%$ of the students dislike peanut butter.



I could represent this in a Bernoulli distribution as



Mean, variance, and standard deviation

Mean

Finding the mean of a Bernoulli random variable is a little counter-intuitive. It seems like we have discrete categories of “dislike peanut butter” and “like peanut butter,” and it doesn’t make much sense to try to find a mean and get a “number” that’s somewhere “in the middle” and means “somewhat likes peanut butter?” It’s all just a little bizarre.

How do we get around this? Well, we mentioned it before, but we assign a value of 0 to the failure category of “dislike peanut butter,” and a value of



1 to the success category of “like peanut butter.” Then we can take the probability weighted sum of the values in our Bernoulli distribution.

$$\mu = (\text{percentage of failures})(0) + (\text{percentage of successes})(1)$$

$$\mu = (0.25)(0) + (0.75)(1)$$

$$\mu = 0 + 0.75$$

$$\mu = 0.75$$

This is the mean of the Bernoulli distribution. Notice how the value we found for the mean is equal to the percentage of “successes.” We said that “liking peanut butter” was a “success,” and then we found that 75 % of our class liked peanut butter, so the mean of the distribution was going to be $\mu = 0.75$.

If we want to create a general formula for finding the mean of a Bernoulli random variable, we could call the probability of success p , and then call the probability of failure $1 - p$ (since total probability always sums to 1, and $p + (1 - p) = p + 1 - p = 1$). Then with failure represented by 0 and success represented by 1, the mean (also called the expected value) will always be

$$\mu = (1 - p)(0) + (p)(1)$$

$$\mu = 0 + p$$

$$\mu = p$$

And we see again that the mean is the same as the probability of success, p . Realize too that, even though we found a mean of $\mu = 0.75$, the



distribution is still discrete. No one in the population is going to take on a value of $\mu = 0.75$; everyone will either be exactly a 0 or exactly a 1.

Variance and standard deviation

We'll use a similar weighting technique to calculate the variance for a Bernoulli random variable. We'll find the difference between both 0 and the mean and 1 and the mean, square that distance, and then multiply by the "weight."

$$\sigma^2 = (0.25)(0 - \mu)^2 + (0.75)(1 - \mu)^2$$

$$\sigma^2 = (0.25)(0 - 0.75)^2 + (0.75)(1 - 0.75)^2$$

$$\sigma^2 = (0.25)(-0.75)^2 + (0.75)(0.25)^2$$

$$\sigma^2 = (0.25)(0.5625) + (0.75)(0.0625)$$

$$\sigma^2 = 0.140625 + 0.046875$$

$$\sigma^2 = 0.1875$$

The standard deviation of a Bernoulli random variable is still just the square root of the variance, so the standard deviation is

$$\sqrt{\sigma^2} = \sqrt{0.1875}$$

$$\sigma = 0.4330$$

The general formula for variance is always given by

$$\sigma^2 = (1 - p)(0 - p)^2 + (p)(1 - p)^2$$

$$\sigma^2 = (1 - p)(-p)^2 + (p)(1 - p)^2$$



$$\sigma^2 = (1 - p)(p^2) + (p)(1 - 2p + p^2)$$

$$\sigma^2 = p^2 - p^3 + p - 2p^2 + p^3$$

$$\sigma^2 = p^2 + p - 2p^2$$

$$\sigma^2 = p - p^2$$

$$\sigma^2 = p(1 - p)$$

Notice that this is just the probability of success p multiplied by the probability of failure $1 - p$. Therefore, standard deviation of the Bernoulli random variable is always given by

$$\sqrt{\sigma^2} = \sqrt{p(1 - p)}$$

$$\sigma = \sqrt{p(1 - p)}$$

