3.5: Variance of Discrete Random Variables

We now look at our second numerical characteristic associated to random variables.

Definition 3.5.1

The ***variance*** of a random variable X is given by

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2],$$

where μ denotes the expected value of X. The ***standard deviation*** of X is given by

$$\sigma = \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

In words, the variance of a random variable is the average of the squared deviations of the random variable from its mean (expected value). Notice that the variance of a random variable will result in a number with units squared, but the standard deviation will have the same units as the random variable. Thus, the standard deviation is easier to interpret, which is why we make a point to define it.

The variance and standard deviation give us a ***measure of spread*** for random variables. The standard deviation is interpreted as a measure of how "spread out" the possible values of X are with respect to the mean of X, $\mu = \mathrm{E}[X]$.

Example 3.5.1

Consider the two random variables X_1 and X_2, whose probability mass functions are given by the histograms in Figure 1 below. Note that X_1 and X_2 have the same mean. However, in looking at the histograms, we see that the possible values of X_2 are more "spread out" from the mean, indicating that the variance (and standard deviation) of X_2 is larger.

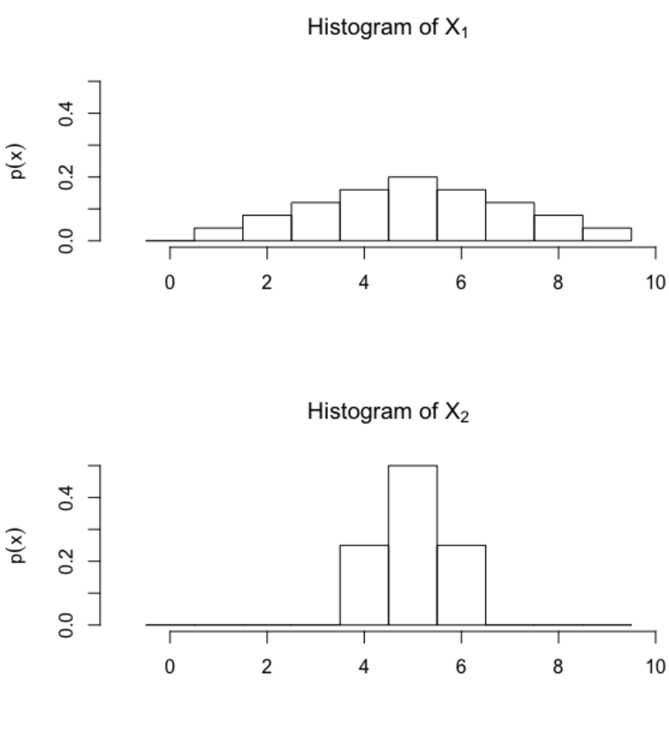

Figure 1: Histograms for random variables X_1 and X_2, both with same expected value different variance.

Theorem 3.4.1 actually tells us how to compute variance, since it is given by finding the expected value of a *function* applied to the random variable. First, if X is a discrete random variable with possible values $x_1, x_2, \ldots, x_i, \ldots$, and probability mass function $p(x)$, then the variance of X is given by

$$\operatorname{Var}(X)=\sum_i (x_i-\mu)^2\cdot p(x_i).$$

The above formula follows directly from Definition 3.5.1. However, there is an alternate formula for calculating variance, given by the following theorem, that is often easier to use.

Theorem 3.5.1

Let X be any random variable, with mean μ. Then the variance of X is

$$\operatorname{Var}(X)=\mathrm{E}[X^2]-\mu^2. \tag{3.5.1}$$

Proof

By the definition of *variance* (Definition 3.5.1) and the linearity of expectation, we have the following:

$$\begin{aligned}\operatorname{Var}(X) &= \mathrm{E}[(X-\mu)^2]\\ &= \mathrm{E}[X^2+\mu^2-2X\mu]\\ &= \mathrm{E}[X^2]+\mathrm{E}[\mu^2]-\mathrm{E}[2X\mu]\\ &= \mathrm{E}[X^2]+\mu^2-2\mu\mathrm{E}[X] \quad \text{(Note: since } \mu \text{ is constant, we can take it out from the expected value)}\\ &= \mathrm{E}[X^2]+\mu^2-2\mu^2\\ &= \mathrm{E}[X^2]-\mu^2\end{aligned}$$

Example 3.5.2

Continuing in the context of Example 3.4.1, we calculate the variance and standard deviation of the random variable X denoting the number of heads obtained in two tosses of a fair coin. Using the alternate formula for variance, we need to first calculate $E[X^2]$, for which we use Theorem 3.4.1:

$$E[X^2]=0^2\cdot p(0)+1^2\cdot p(1)+2^2\cdot p(2)=0+0.5+1=1.5.$$

In Example 3.4.1, we found that $\mu=E[X]=1$. Thus, we find

$$\begin{aligned}\operatorname{Var}(X) &= E[X^2]-\mu^2=1.5-1=0.5\\ \Rightarrow\ \mathrm{SD}(X) &= \sqrt{\operatorname{Var}(X)}=\sqrt{0.5}\approx 0.707\end{aligned}$$

Exercise 3.5.1

Consider the context of Example 3.4.2, where we defined the random variable X to be our winnings on a single play of game involving flipping a fair coin three times. We found that $\mathrm{E}[X]=1.25$. Now find the variance and standard deviation of X.

Answer

First, find $\mathrm{E}[X^2]$:

$$\begin{aligned}\mathrm{E}[X^2] &= \sum_i x_i^2\cdot p(x_i)\\ &= (-1)^2\cdot\frac{1}{8}+1^2\cdot\frac{1}{2}+2^2\cdot\frac{1}{4}+3^2\cdot\frac{1}{8}=\frac{11}{4}=2.75\end{aligned}$$

Now, we use the alternate formula for calculating variance:

$$\begin{aligned}\operatorname{Var}(X) &= \mathrm{E}[X^2]-\mathrm{E}[X]^2=2.75-1.25^2=1.1875\\ \Rightarrow \mathrm{SD}(X) &= \sqrt{1.1875}\approx 1.0897\end{aligned}$$

Given that the variance of a random variable is defined to be the expected value of *squared* deviations from the mean, variance is not linear as expected value is. We do have the following useful property of variance though.

Theorem $3.5.2$

Let X be a random variable, and a, b be constants. Then the following holds:

$$\operatorname{Var}(aX+b) = a^2\operatorname{Var}(X).$$

Exercise $3.5.2$

Prove Theorem 3.5.2.

Answer

First, let $\mu = \mathrm{E}[X]$ and note that by the linearity of expectation we have

$$\mathrm{E}[aX+b] = a\mathrm{E}[X]+b = a\mu+b.$$

Now, we use the alternate formula for variance given in Theorem 3.5.1 to prove the result:

$$\begin{aligned}\operatorname{Var}(aX+b) &= \mathrm{E}\left[(aX+b)^2\right] - \left(\mathrm{E}[aX+b]\right)^2 \\ &= \mathrm{E}[a^2X^2+2abX+b^2] - (a\mu+b)^2 \\ &= a^2\mathrm{E}[X^2]+2ab\mathrm{E}[X]+b^2-a^2\mu^2-2ab\mu-b^2 \\ &= a^2\mathrm{E}[X^2]-a^2\mu^2 = a^2(\mathrm{E}[X^2]-\mu^2) = a^2\operatorname{Var}(X)\end{aligned}$$

Theorem 3.5.2 easily follows from a little algebraic modification. Note that the "$+b$" disappears in the formula. There is an intuitive reason for this. Namely, the "$+b$" corresponds to a *horizontal shift* of the probability mass function for the random variable. Such a transformation to this function is not going to affect the *spread*, i.e., the variance will not change.

As with expected values, for many of the common probability distributions, the variance is given by a parameter or a function of the parameters for the distribution.

Variance for Discrete Distributions

Distribution	**Expected Value**
Bernoulli(p)	$p(1-p)$
binomial(n,p)	$np(1-p)$
hypergeometric(N,n,m)	$\frac{n(m/N)(1-m/N)(N-n)}{N-1}$
geometric(p)	$\frac{1-p}{p^2}$
negative binomial(r,p)	$\frac{r(1-p)}{p^2}$
Poisson(λ)	λ

This page titled 3.5: Variance of Discrete Random Variables is shared under a not declared license and was authored, remixed, and/or curated by Kristin Kuter.