Chapter 7.2

Variance of a Discrete Random Variable

Review of Last Week

Binomial random variable: counts the number of successes in n **independent** trials with **constant** probability of success equal to π on each trial.

$$X \sim Binom(n, \pi)$$

$$f(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

Binomial calculations in R:

- `dbinom(x, size, prob)` calculates $P(X = x)$
- `pbinom(q, size, prob)` calculates $P(X \leq q)$
- `pbinom(q, size, prob, lower.tail = F)` calculates $P(X > q)$

Review of Last Week

Expected Value average value of the random variable across replications of the random experiment.

- $E[X] = \sum_x xf(x)$
- $E[X] = n\pi$ when $X \sim Binom(n, \pi)$.
- Linearity of expected value:

$$E[aX + b] = aE[X] + b.$$

- Law of the unconscious probabilist:

$$E[t(X)] = \sum_x t(x)f(x).$$

Warm up

Let's consider two random variables X and Y with PMF as follows. Both have an expected value of 3. But which will deviate more from expected?

x	1	2	3	4	5
f_X	0.1	0.2	0.4	0.2	0.1
f_Y	0.3	0.1	0.2	0.1	0.3

Variance of a random variable

Definition 7.2 Let X be a discrete random variable with mean μ. The **variance** of X, assuming it exists, is defined by

$$\sigma_X^2 = Var\left[X\right] = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 \times f(x).$$

The **standard deviation** is the positive square root of the variance.

- The interpretation attached to the variance is that small values imply that X is very likely to be close to $E[X]$ and larger values imply X is more variable.
- The standard deviation has this same qualitative interpretation but is easier to interpret in that the unit of measurement is the same as that for the original random variable. (The unit of measurement on the variance is the square of the original unit.)

An alternate formula for $Var[X]$ that is not as intuitive as the definition, but is often easier to use in practice.

$$
\begin{aligned}
Var[X] &= E\left[(X-\mu)^2\right], \\
&= \sum_x (x-\mu)^2 \cdot f(x), \\
&= \sum_x (x^2 - 2x\mu + \mu^2) \cdot f(x), \\
&= \sum_x x^2 \cdot f(x) - \sum_x 2x\mu f(x) + \sum_x \mu^2 f(x) \\
&= E\left[X^2\right] - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x), \\
&= E\left[X^2\right] - 2\mu \cdot \mu + \mu^2 \cdot 1, \\
&= E\left[X^2\right] - \mu^2.
\end{aligned}
$$

Example 7.6

Calculate $Var[X]$ if X represents the outcome when a fair die is rolled.

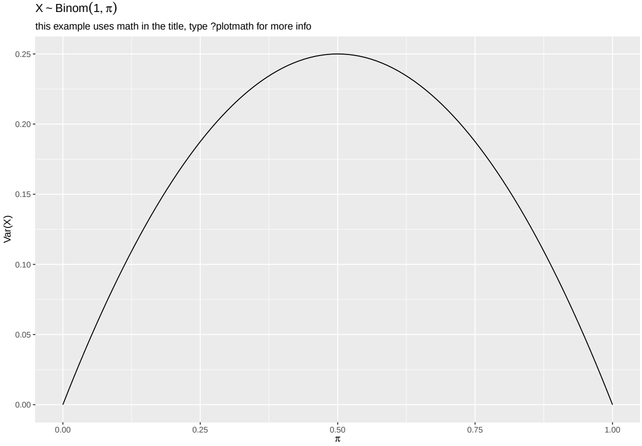
Example 7.6 contd.

Calculate $Var[X]$ if X represents the outcome when a fair die is rolled.

Example 7.7

If $X \sim Binom(1, \pi)$, then X is called a **Bernoulli** random variable. For a Bernoulli random variable, what are $E[X]$ and $Var[X]$?

As a function of π, $Var(X)$ in example 7.7 is quadratic. Its graph is a parabola that opens downward and the largest variance occurs when $\pi = \frac{1}{2}$.



```
ggplot() +
  geom_function(fun = function(x){x*(1-x)}, xlim = c(0,1) ) +
  labs(x = expression(pi),
       y = "Var(X)",
       title = expression(X %~% Binom(1, pi)),
       subtitle = "type ?plotmath for more info" )
```

Variance of a Linear Transformation

Lemma 7.4 Let X be a discrete random variable and suppose a and b are constants. Then

$$Var(aX + b) \quad = \quad a^2 Var(X).$$

Example 7.8

Suppose $E(X) = \mu$ and the standard deviation of X equals σ. Find the mean and variance of the new random variable

$$Y = (X - \mu)/\sigma.$$

The fact that

$$Y = (X - \mu)/\sigma$$

has mean 0 and standard deviation 1 is the reason it is often referred to as the *standardization* of X.

Chebyshev's Inequality

Lemma 7.5 For any random variable X with mean μ and standard deviation σ:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2},$$

where k is some positive number.

Note: $|X - \mu| \geq k\sigma$ means either

$$(X - \mu) \geq k\sigma \Rightarrow X \geq \mu + k\sigma$$

or

$$(X - \mu) \leq -k\sigma \Rightarrow X \leq \mu - k\sigma.$$

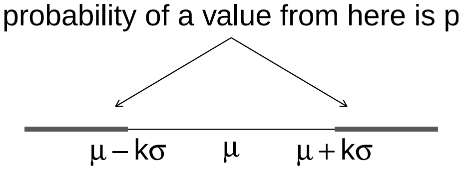
That is, Chebyshev's inequality guarantees that at most $\frac{1}{k^2}$ of the distribution can be k or more standard deviations from the mean

Chebyshev's Inequality: heuristic proof

Let us denote

$$p = P(|X - \mu| \geq k\sigma). \qquad (1)$$

For any X which satisfies equation (1)



Chebyshev (heuristic proof)

One PMF which satisfies equation (1) is the following:

value	$\mu - k\sigma$	μ	$\mu + k\sigma$
probability	$\frac{p}{2}$	$1 - p$	$\frac{p}{2}$

This PMF yields the smallest variance that is possible for any distribution that satisfies 1.

Chebyshev's Inequality: heuristic proof

It is straightforward to check that the variance of this distribution

value	$\mu - k\sigma$	μ	$\mu + k\sigma$
probability	$\frac{p}{2}$	$1-p$	$\frac{p}{2}$

is $pk^2\sigma^2$ and therefore

$$Var(X) = \sigma^2 \geq pk^2\sigma^2$$

from which it follows that

$$p = P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Example 7.9

What does Chebyshev's inequality say about the probability that a random variable is more than 2 standard deviations from its mean? Calculate the exact probability for the outcome, X, when a fair die is rolled.

Chebyshev's inequality only provides informative answers for values of k greater than 1. In addition, the bound $\frac{1}{k^2}$ is usually much larger than the exact probability for many distributions.