Some Properties of Expected Value & Variance

Definitions and Notation

Let's suppose that x and y are discrete variables with a joint mass function $f(x, y)$. The marginal mass functions are defined by

$$\begin{aligned} f_1(x) &= \sum_y f(x, y) \\ f_2(y) &= \sum_x f(x, y) \end{aligned}$$

To save on subscripts, let's let $p(x) = f_1(x)$ and $q(y) = f_2(x)$.

We will say that x and y are **independent** if $f(x, y) = p(x) \cdot q(y)$. Independence is required for some of the statements below to be true.

Properties of Means and Variances

The following can be verified by looking at the sums involved and using some algebra:

Mean of a Sum

For any discrete variables x and y, $\mu_{x+y} = \mu_x + \mu_y$. (Note: independence not required.)

$$\begin{aligned} \mu_{x+y} &= \sum_{x,y} (x + y) f(x, y) \\ &= \sum_{x,y} x \cdot f(x, y) + y \cdot f(x, y) \\ &= \sum_x \sum_y x \cdot f(x, y) + \sum_y \sum_x y \cdot f(x, y) \\ &= \sum_x x \sum_y f(x, y) + \sum_y y \sum_x f(x, y) \\ &= \sum_x x \cdot p(x) + \sum_y y \cdot q(y) \\ &= \mu_x + \mu_y \end{aligned}$$

Mean of an Independent Product

If x and y are independent, then $\mu_{xy} = \mu_x \cdot \mu_y$.

$$\begin{aligned}
\mu_{xy} &= \sum_{x,y} (xy) f(x,y) \\
&= \sum_{x,y} (xy) p(x) q(y) \\
&= \sum_{x} \sum_{y} xy \cdot p(x) q(y) \\
&= \sum_{x} x p(x) \sum_{y} y \cdot q(y) \\
&= \sum_{x} x p(x) \cdot \mu_y \\
&= \mu_x \cdot \mu_y
\end{aligned}$$

Another Formula for Variance

$V(x) = \mu_{x^2} - (\mu_x)^2$

This was Exercise 2.25.

Using this, we can prove a result about the variance of the sum of independent variables. This is easier to express using a new notation. Let's use $V(x)$ for the variance of x, and $E(x)$ for the mean of x. (E comes from *expected value*).

Variance of an Independent Sum

If x and y are independent, then $V(x+y) = V(x) + V(y)$.

$$\begin{aligned}
V(x+y) &= E((x+y)^2) - [E(x+y)]^2 \\
&= E(x^2 + 2xy + y^2) - [(E(x))^2 + 2E(x)E(y) + (E(y))^2] \\
&= E(x^2) + E(2xy) + E(y^2) - [(E(x))^2 + 2E(x)E(y) + (E(y))^2] \\
&= E(x^2) + 2E(x)E(y) + E(y^2) - [E(x)^2 + 2E(x)E(y) + E(y)^2] \\
&= E(x^2) + E(y^2) - E(x)^2 - E(y)^2 \\
&= E(x^2) - E(x)^2 + E(y^2) - E(y)^2 \\
&= V(x) + V(y)
\end{aligned}$$

Application to the Binomial Distribution

If $x \sim \text{Bin}(n, \pi)$, we can use these properties to determine the mean and variance of x by writing

$$x = x_1 + x_2 + \cdots x_n ,$$

where each $x_i \sim \text{Bin}(1, \pi)$ and represents the outcome of a single trial. Note that the value of x_i is either 0 (failure) or 1 (success), and π is the proportion of the time that there is success.

$$\begin{aligned} E(x) &= E(x_1 + x_2 + \cdots + x_n) \\ &= E(x_1) + E(x_2) + \cdots + E(x_n) \\ &= \pi + \pi + \cdots + \pi = n\pi \end{aligned}$$

$$\begin{aligned} V(x) &= V(x_1 + x_2 + \cdots + x_n) \\ &= V(x_1) + V(x_2) + \cdots + V(x_n) \\ &= \pi(1-\pi) + \pi(1-\pi) + \cdots + \pi(1-\pi) \\ &= n\pi(1-\pi) \end{aligned}$$

Exercises

Let x and y be independent variables with mean and standard deviation as shown in the table below:

	x	y
mean	10	5
standard deviation	3	2

Determine the mean and standard deviation of the following: (a) $x + y$, (b) $x - y$, (c) xy (d) $4x + 6y$