1 Variance

1. **(original)** Which of the following situations are possible? Explain.

   (a) A random variable X_1 has expected value 1 and variance $\frac{1}{2}$.

   This is possible. Consider a binomial random variable with $n = 2$ and $p = \frac{1}{2}$. Its expected value is $np = 2 \cdot \frac{1}{2} = 1$ and its variance is $np(1-p) = 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$.

   (b) A random variable X_2 has expected value 1 and variance $-\frac{1}{2}$.

   This is impossible. Variance must always be nonnegative.

   (c) A random variable X_3 has expected value 1 and variance 1000000.

   This is possible. Consider a random variable with $P(X = 1001) = \frac{1}{2}$ and $P(X = -999) = \frac{1}{2}$. We have $E[X] = \frac{1}{2} \cdot 1001 + \frac{1}{2} \cdot (-999) = 1$. We have $\text{Var}[X] = E[(X - E[X])^2] = \frac{1}{2}(1001 - 1)^2 + \frac{1}{2}(-999 - 1)^2 = 1000000$.

   (d) A random variable X_4 has expected value 1, variance 4, and standard error 4.

   This is impossible. The standard error must be equal to the square root of the variance.

   (e) A Poisson random variable X_5 has expected value 1 and variance 2.

   This is impossible. For a Poisson random variable with parameter λ, the expected value and the variance are both equal to λ, so a random variable with expected value 1 and variance 2 cannot be a Poisson random variable.

   (f) Two inepdendent random variables X and Y have $E[X] = 1$, $E[Y] = 1$, $\text{Var}[X] = 2$, and $\text{Var}[Y] = 3$, and their sum $X + Y$ has $E[X + Y] = 2$ and $\text{Var}[X + Y] = 4$.

   This is impossible. For two *independent* random variables X and Y, we must have that $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$, which is not the case in this putative example.

2. **(original)** I flip a fair coin twice. Let X denote the number of heads and Y denote the number of tails.

   (a) What is $\text{Var}[X]$?

   We have a binomial random variable with $n = 2$ and $p = \frac{1}{2}$, so the variance is $np(1-p) = \boxed{\frac{1}{2}}$.

(b) What is $\text{Var}[Y]$?

This is the exact same situation as part (a), where we consider tails to be a success instead of heads, so the variance is $\boxed{\frac{1}{2}}$.

(c) What is $\text{Var}[X+Y]$?

These two random variables are *not* independent, so we don't just sum our answers for part (a) and part (b). Instead, we need to use one of the formulas for variance. We can use $\text{Var}[X+Y] = E[(X+Y)^2] - (E[X+Y])^2$.

There are four possible outcomes (HH, HT, TH, TT). In each case, the sum of the number of heads and number of tails is 2, which makes sense since each flip is either heads or tails. We've shown that $X+Y$ is always equal to 2, i.e. that the random variable $X+Y$ is equal to 2 with probability 1. We have $E[X+Y]=2$ and similarly that $E[(X+Y)^2]=4$. We thus have $\text{Var}[X+Y] = 4-2^2 = 0$.

3. **(inspired by pset 19 #1)** A binomial random variable X has mean 20 and variance 4. Can you find n and p for this random variable? What is $P(X=2)$?

We know that the mean for a binomial random variable is np and the variance is $np(1-p)$. Knowing the mean and variance, we can set up a system of equations

$$np = 20$$
$$np(1-p) = 4.$$

Dividing the second equation by the first, we see that $1-p=\frac{1}{5}$, so $p=\frac{4}{5}$. Plugging this into the first equation, we see that $n=20$. We then have

$$P(X=2) = \boxed{\binom{25}{2}\left(\frac{4}{5}\right)^2\left(\frac{1}{5}\right)^{23}}.$$

4. **(original)** Suppose playing a gambling game at one casino has an expected value of -1 dollars and a variance of 50 cents^2. At a pricier casino across the street, the exact same game (with the same probabilities of various levels of payouts) is played except the amount of money you pay to play is 10 times as high and all payouts are 10 times as high as well. What is the expected value and variance of this game at the pricier casino?

Let X denote the random variable describing your winnings at the game at the cheaper casino. At the more expensive casino, all prices and payouts are 10 times as high, so the more expensive casino is described by the random variable $10X$. $E[10X] = 10E[X]$, so the expected value is $\boxed{-10}$ dollars, and $\text{Var}[10X] = 100\,\text{Var}[X]$, so the variance is $\boxed{50}$ dollars^2.

2 Covariance and limits of random variables

5. **(based on pset 20 #1)** Let X_1, X_2, and X_3 denote the numbers that come up on three rolls of a fair *four-sided* die. Let $X = X_1 + X_2 + X_3$, $Y = X_1 + X_2$, $\overline{X} = X/3$, $\mu = E[X_1]$, and $\sigma^2 = \text{Var}[X_1]$.

   (a) What are the ranges of X_1, X_2, X_3, X, $\overline{X}$, and Y?

   X_1, X_2, and X_3 all have range $\{1, 2, 3, 4\}$. X has range $\{3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, all possible sums of three numbers from 1 to 4. $\overline{X}$ has range $\{1, \frac{3}{4}, \ldots, \frac{11}{3}, 4\}$, all possible fractions with an element of the range of X in the numerator and 3 in the denominator. Y has range $\{2, 3, 4, 5, 6, 7, 8\}$, all possible sums of two numbers from 1 to 4.

   (b) Find the variances of X_1, X_2, X_3, X, $\overline{X}$, and Y.

   For a single four-sided die, we compute

$$\begin{aligned} E[X_1] &= \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 3 + \frac{1}{4} \cdot 4 \\ &= \frac{5}{2} \\ E[X_1^2] &= \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 4 + \frac{1}{4} \cdot 9 + \frac{1}{4} \cdot 16 \\ &= \frac{15}{2}. \end{aligned}$$

   The variance of X_1 is thus $E[X_1^2] - E[X_1]^2 = \frac{15}{2} - \frac{25}{4} = \frac{5}{4}$.
   Since X_2 and X_3 are independent of and identically distributed to X_1, their variances are also $\frac{5}{4}$.
   Since $X = X_1 + X_2 + X_3$ and X_1, X_2, X_3 are independent, we have

$$\text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \text{Var}[X_3] = \frac{15}{4}.$$

   Similarly, we have

$$\text{Var}[\overline{X}] = \text{Var}\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{1}{9}\text{Var}[X_1 + X_2 + X_3] = \frac{1}{9} \cdot \frac{15}{4} = \frac{5}{12}.$$

   Finally, since $Y = X_1 + X_2$ and X_1 and X_2 are independent, we have $E[Y] = \text{Var}[X_1] + \text{Var}[X_2] = \frac{5}{2}$.

   (c) Find $\text{Cov}[X_1, Y]$ and $\text{Cov}[X_3, Y]$.

   We have, using the definition of covariance, the fact that the expected value of a sum is equal to the sum of the expected values, and the fact that the expected value of a product of two independent random variables is the product of their expected values,

$$\begin{aligned}\text{Cov}[X_1, Y] &= E[X_1(X_1 + X_2)] - E[X_1]E[X_1 + X_2] \\ &= E[X_1^2 + X_1X_2] - E[X_1]\left(E[X_1] + E[X_2]\right) \\ &= E[X_1^2] + E[X_1X_2] - E[X_1]^2 - E[X_1]E[X_2] \\ &= E[X_1^2] + E[X_1]E[X_2] - E[X_1]^2 - E[X_1]E[X_2] \\ &= E[X_1^2] - E[X_1]^2 \\ &= \frac{5}{4}.\end{aligned}$$

X_3 and Y are independent (as Y is built out of X_1 and X_2, both of which are independent of X_3), so the covariance $\text{Cov}[X_3, Y] = 0$. Alternatively, we could carry out a computation like the one above.

(d) Are X_1 and Y independent? Why or why not? What about X_3 and Y?

Since we know $\text{Cov}[X_1, Y] \neq 0$, X_1 and Y are not independent. X_3 and Y, however, are.

(e) Let $\overline{\mu} = E[\overline{X}]$ and $\overline{\sigma}^2 = \text{Var}[\overline{X}]$. If $Z = \frac{\overline{X} - \overline{\mu}}{\overline{\sigma}}$, find $E[Z]$ and $\text{Var}[Z]$.

We have

$$\begin{aligned}E[Z] &= E\left[\frac{\overline{X} - \overline{\mu}}{\overline{\sigma}}\right] \\ &= \frac{1}{\overline{\sigma}}E[\overline{X} - \overline{\mu}] \\ &= \frac{1}{\overline{\sigma}}(E[\overline{X}] - \overline{\mu}]) \\ &= 0\end{aligned}$$

and

$$\begin{aligned}\text{Var}[Z] &= \text{Var}\left[\frac{\overline{X} - \overline{\mu}}{\overline{\sigma}}\right] \\ &= \frac{1}{\overline{\sigma}^2}\,\text{Var}[\overline{X} - \overline{\mu}].\end{aligned}$$

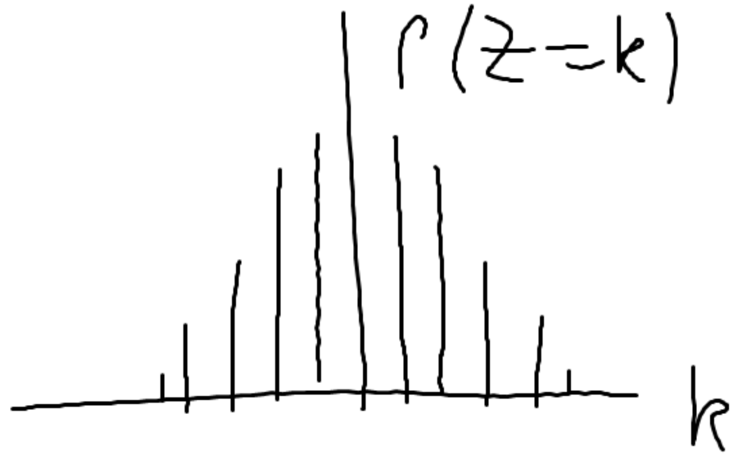
We know that $\overline{X}$ and $\overline{X} - \overline{\mu}$ have the same variance (for example, because the constant random variable $\overline{\mu}$ is independent of $\overline{X}$), so $\text{Var}[\overline{X} - \overline{\mu}] = \text{Var}[X] = \sigma^2$, and thus $\text{Var}[Z] = 1$.

(f) What is the pmf of Z? Sketch a picture.

The random variable Z takes on the following values: $\frac{-3/2}{\sqrt{5/12}}$, $\frac{-7/6}{\sqrt{5/12}}$, $\frac{-5/6}{\sqrt{5/12}}$, $\frac{-1/2}{\sqrt{5/12}}$, $\frac{-1/6}{\sqrt{5/12}}$, $\frac{1/6}{\sqrt{5/12}}$, $\frac{1/2}{\sqrt{5/12}}$, $\frac{5/6}{\sqrt{5/12}}$, $\frac{7/6}{\sqrt{5/12}}$, and $\frac{3/2}{\sqrt{5/12}}$. To find the probability that Z takes on one of these values, we can note that they correspond to the values 3, 4, and so on through 12 of the random variable X, then count the number of ways to get those integers as the sum of three dice rolls. For example, $P\left(\frac{1/6}{\sqrt{5/12}}\right) = \frac{12}{64} = \frac{3}{16}$ because an 8 can be rolled in the following 12 ways: $1+3+4$ (which can be arranged in 6 ways), $2+2+4$ (which can be arranged in 3 ways), and $2+3+3$ (which can be arranged in 3 ways). We have that the probabilities are, in order corresponding to the order of the range given above,

$$\frac{1}{64}, \frac{3}{64}, \frac{3}{32}, \frac{5}{32}, \frac{3}{16}, \frac{3}{16}, \frac{5}{32}, \frac{3}{32}, \frac{3}{64}, \frac{1}{64}.$$

A rough sketch of the pmf is below:



All problems labeled "inspired by pset" adapted from the problem sets. Problem set 19 is adapted from previous 10B HWs written by Anna Seigal. Problem set 20 is new for this year's iteration math 10B.