Probability: Random Variables

The Expectation and Variance of the Sample Mean

Suppose we roll a fair die 100 times and get the sample average, then another 100 times and get the average of that 100 rolls and repeat this many times. Each time the sample average would be different because it is random and the probability density function for the distribution of these sample averages will look like this:

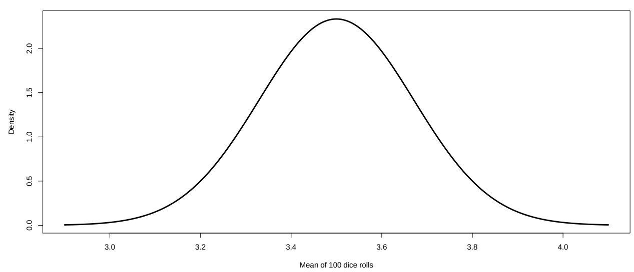

Figure 1: Distribution of the mean of 100 Dice Rolls

In this document we show what is the expected value and variance of this probability distribution.

First we introduce two important rules about expected values:

Rule 1: $E(a + bX) = a + bE(X)$ where a and b are any numbers
Rule 2: $E(X + Y) = E(X) + E(Y)$

The mean value of 100 dice rolls is

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_{100}}{100}$$

The expected value of distribution of the mean is

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \cdots + X_{100}}{100}\right) = E\left(\frac{X_1}{100} + \frac{X_2}{100} + \cdots + \frac{X_{100}}{100}\right)$$

Using Rule 2

$$= E\left(\frac{X_1}{100}\right) + E\left(\frac{X_2}{100}\right) + \cdots + E\left(\frac{X_{100}}{100}\right)$$

Using Rule 1, with $a = 0$ and $b = \frac{1}{100}$

$$= \frac{1}{100}E(X_1) + \frac{1}{100}E(X_2) + \cdots + \frac{1}{100}E(X_{100}) = \frac{1}{100}[E(X_1) + E(X_2) + \cdots + E(X_{100})]$$

Since $E(X) = 3.5$ we get:

$$= \frac{1}{100}[3.5 + 3.5 + \cdots + 3.5] = \frac{1}{100} \times 3.5 \times 100 = 3.5 = E(X)$$

Hence the expectation of sample mean is the same as the expectation of original random variable X and is equal to 3.5.

Next let's look at the variance of the mean distribution.
We start with two rules about variances:

Rule 1: $Var(a + bX) = b^2 Var(X)$ where a and b are any numbers
Rule 2: $Var(X + Y) = Var(X) + Var(Y)$, if X and Y are independent

Now we can calculate the variance of the distribution of means:

$$Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \cdots + X_{100}}{100}\right)$$

Using Rule 1 with $a = 0$ and $b = \frac{1}{100}$

$$= \frac{1}{100^2}Var(X_1 + X_2 + \cdots + X_{100})$$

Using Rule 2, since $X_1, X_2, , X_{100}$ are independent

$$= \frac{1}{100^2}[Var(X_1) + Var(X_2) + \cdots + Var(X_{100})]$$

Since $Var(X) = 2.92$ we get

$$= \frac{1}{100^2}[2.92 + 2.92 + \cdots + 2.92] = \frac{1}{100^2} \times 2.92 \times 100 = \frac{2.92}{100} = \frac{Var(X)}{100}$$

So the variance of the mean is one hundredth of the variance of original random variable X.

Here is the summary of the results:
$E(\bar{X}) = 3.5$
$Var(\bar{X}) = \frac{2.92}{100}$
$SD(\bar{X}) = \sqrt{\frac{2.92}{100}} = 0.17$

To picture these quantities on the graph of distribution of means, we put a vertical bar at 3.5 for the expectation (red line) and two vertical (green) lines at 3.33 ($E(\bar{X}) - SD(\bar{X}) =$

$3.5 - 0.17$) and at 3.67 ($E(\bar{X}) + SD(\bar{X}) = 3.5 + 0.17$) which represent expectation plus one standard deviation and expectation minus one standard deviation.

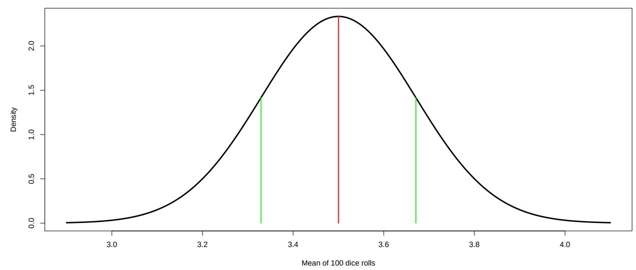

Figure 2: Distribution of the mean of 100 Dice Rolls with expectation and one standard deviation