1 Variance

(a) Let X be a random variable representing the outcome of the roll of one fair 6-sided die. What is $\text{Var}(X)$?

(b) Let Z be a random variable representing the average of n rolls of a fair die 6-sided die. What is $\text{Var}(Z)$?

(c) A building has n floors numbered $1, 2, \dots, n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the *variance* of the number of floors the elevator *does not* stop at? (In fact, the variance of the number of floors the elevator *does* stop at must be the same (make sure you understand why), but the former is a little easier to compute.)

Solution:

(a) Recall that $\text{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$. We can compute each of the individual terms using the definition of expectation:

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2} \\
\mathbb{E}\left[X^2\right] &= \frac{1}{6}\left(1^2+2^2+3^2+4^2+5^2+6^2\right) \\
&= \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}
\end{aligned}$$

Now, we plug back into the variance expression:

$$\begin{aligned}
\text{Var}(X) &= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 \\
&= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}
\end{aligned}$$

(b) Because each dice roll is independent of the others, we can utilize the fact that for independent random variables X and Y, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$. Let X_i be a random variable representing the outcome of the ith dice roll. We now have:

$$\begin{aligned}
\mathrm{Var}(Z) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \left(\frac{1}{n}\right)^2 \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathrm{Var}(X_i) && \text{All } X_i\text{'s are independent.} \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \frac{35}{12} && \text{We computed the variance of one roll in part (a).} \\
&= \left(\frac{1}{n}\right)^2 \cdot n \cdot \frac{35}{12} = \frac{35}{12n}
\end{aligned}$$

(c) Let X be the number of floors the elevator does not stop at. We can represent X as the sum of the indicator variables $X_1, \ldots, X_n$, where $X_i = 1$ if no one gets off on floor i. Thus, we have

$$\mathbb{E}[X_i] = \mathbb{P}[X_i = 1] = \left(\frac{n-1}{n}\right)^m,$$

and from linearity of expectation,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\left(\frac{n-1}{n}\right)^m.$$

To find the variance, we cannot simply sum the variance of our indicator variables. However, we can still compute $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ directly using linearity of expectation, but now how can we find $\mathbb{E}[X^2]$? Recall that

$$\mathbb{E}[X^2] = \mathbb{E}[(X_1 + \cdots + X_n)^2] = \mathbb{E}\left[\sum_{i,j} X_i X_j\right] = \sum_{i,j} \mathbb{E}[X_i X_j] = \sum_{i}^{n} \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i X_j].$$

The first term is simple to calculate - Note that the squared expectation of an indicator is still just $\mathbb{P}[X = 1]$.

$$\mathbb{E}[X_i^2] = 1^2 \mathbb{P}[X_i = 1] = \left(\frac{n-1}{n}\right)^m,$$

There are n terms in our summation. Thus,

$$\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] = n\left(\frac{n-1}{n}\right)^m.$$

Next, $X_i X_j = 1$ when both X_i and X_j are 1, which means no one gets off the elevator on floor i and floor j. This happens with probability

$$\mathbb{P}[X_i = X_j = 1] = \mathbb{P}[X_i = 1 \cap X_j = 1] = \left(\frac{n-2}{n}\right)^m.$$

There are n(n-1) terms in our summation. You could count this with order, directly seeing that there are n options for i and then n-1 options for j. Or, unordered you can see $\binom{n}{2}$, then multiply by 2 since each X_iX_j term shows up twice. Verify for yourself why this is the case. (How many cross terms are in $(x_1+x_2+x_3)^2$?) Thus,

$$\sum_{i\neq j}\mathbb{E}[X_iX_j] = n(n-1)\left(\frac{n-2}{n}\right)^m.$$

Finally, we plug in to see that

$$\operatorname{Var}(X) = \mathbb{E}[X^2]-\mathbb{E}[X]^2 = n\left(\frac{n-1}{n}\right)^m + n(n-1)\left(\frac{n-2}{n}\right)^m - n^2\left(\frac{n-1}{n}\right)^{2m}.$$

2 Inequality Practice

(a) X is a random variable such that $X > -5$ and $\mathbb{E}[X] = -3$. Find an upper bound for the probability of X being greater than or equal to -1.

(b) Y is a random variable such that $Y < 10$ and $\mathbb{E}[Y] = 1$. Find an upper bound for the probability of Y being less than or equal to -1.

(c) You roll a die 100 times. Let Z be the sum of the numbers that appear on the die throughout the 100 rolls. Compute $\operatorname{Var}(Z)$. Then use Chebyshev's inequality to bound the probability of the sum Z being greater than 400 or less than 300.

Solution:

(a) We want to use Markov's Inequality, but recall that Markov's Inequality only works with non-negative random variables. So, we define a new random variable $\tilde{X} = X+5$, where $\tilde{X}$ is always non-negative, so we can use Markov's on $\tilde{X}$. By linearity of expectation, $\mathbb{E}[\tilde{X}] = -3+5 = 2$. So, $\mathbb{P}[\tilde{X}\geq 4]\leq 2/4 = 1/2$.

(b) We again use Markov's Inequality. Similarly, define $\tilde{Y} = -Y+10$, and $\mathbb{E}[\tilde{Y}] = -1+10 = 9$. $P[Y\leq -1] = P[-Y\geq 1] = P[-Y+10\geq 11]\leq 9/11$.

(c) Let Z_i be the number on the die for the ith roll, for $i = 1,\ldots,100$. Then, $Z = \sum_{i=1}^{100} Z_i$. By linearity of expectation, $\mathbb{E}[Z] = \sum_{i=1}^{100}\mathbb{E}[Z_i]$.

$$\mathbb{E}[Z_i] = \sum_{j=1}^{6} j\cdot\mathbb{P}[Z_i = j] = \sum_{j=1}^{6} j\cdot\frac{1}{6} = \frac{1}{6}\cdot\sum_{j=1}^{6} j = \frac{1}{6}\cdot 21 = \frac{7}{2}$$

Then, we have $\mathbb{E}[Z] = 100\cdot(7/2) = 350$.

$$\mathbb{E}[Z_i^2] = \sum_{j=1}^{6} j^2\cdot\mathbb{P}[Z_i = j] = \sum_{j=1}^{6} j^2\cdot\frac{1}{6} = \frac{1}{6}\cdot\sum_{j=1}^{6} j^2 = \frac{1}{6}\cdot 91 = \frac{91}{6}$$

Then, we have

$$\text{Var}(Z_i) = \mathbb{E}[Z_i^2] - \mathbb{E}[Z_i]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12},$$

Since the Z_is are independent, and therefore uncorrelated, we can add the $\text{Var}(Z_i)$s to get $\text{Var}(Z) = 100(35/12)$.

Finally, we note that we can upper bound $\mathbb{P}[|Z-350| > 50]$ with $\mathbb{P}[|Z-350| \geq 50]$.

Putting it all together, we use Chebyshev's to get

$$\mathbb{P}[|Z-350| > 50] < \mathbb{P}[|Z-350| \geq 50] \leq \frac{100(35/12)}{50^2} = \frac{7}{60}.$$

3 Working with the Law of Large Numbers

(a) A fair coin is tossed multiple times and you win a prize if there are more than 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(b) A fair coin is tossed multiple times and you win a prize if there are more than 40% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(c) A fair coin is tossed multiple times and you win a prize if there are between 40% and 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(d) A fair coin is tossed multiple times and you win a prize if there are exactly 50% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

Solution:

(a) 10 tosses. By LLN, the sample mean should have higher probability to be close to the population mean as n increases. Therefore the average proportion of coins that are heads should be closer to 0.50, and has a lower chance of being greater than 0.60 if there are 100 tosses (compared with 10 tosses).

(b) 100 tosses. Again, by LLN, the sample mean should have higher probability to be close to the population mean as n increases. Therefore the average proportion of coins that are heads should be closer to 0.50, and has a lower chance of being smaller than 0.40 if there are 100 tosses. A lower chance of being smaller than 0.40 is the desired result.

(c) 100 tosses. Again, by LLN, the average proportion of coins that are heads should be closer to 0.50, and has a lower chance of being both smaller than 0.40 if there are 100 tosses. Similarly, there is a lower chance of being larger than 0.60 if there are 100 tosses. Lower chances of both of these events is desired if we want the fraction of heads to be between 0.4 and 0.6.

(d) 10 tosses. Compare the probability of getting equal number of heads and tails between $2n$ and $2n+2$ tosses.

$$
\begin{aligned}
\mathbb{P}[n \text{ heads in } 2n \text{ tosses}] &= \binom{2n}{n}\frac{1}{2^{2n}} \\
\mathbb{P}[n+1 \text{ heads in } 2n+2 \text{ tosses}] &= \binom{2n+2}{n+1}\frac{1}{2^{2n+2}} = \frac{(2n+2)!}{(n+1)!(n+1)!}\cdot\frac{1}{2^{2n+2}} \\
&= \frac{(2n+2)(2n+1)2n!}{(n+1)(n+1)n!n!}\cdot\frac{1}{2^{2n+2}} \\
&= \frac{2n+2}{n+1}\cdot\frac{2n+1}{n+1}\binom{2n}{n}\cdot\frac{1}{2^{2n+2}} < \left(\frac{2n+2}{n+1}\right)^2\binom{2n}{n}\cdot\frac{1}{2^{2n+2}} \\
&= 4\binom{2n}{n}\cdot\frac{1}{2^{2n+2}} = \binom{2n}{n}\frac{1}{2^{2n}} = \mathbb{P}[n \text{ heads in } 2n \text{ tosses}]
\end{aligned}
$$

As we increment n, the probability will always decrease. Therefore, the larger n is, the less probability we'll get exactly 50% heads. $\square$

Note: By Stirling's approximation, $\binom{2n}{n}2^{-2n}$ is roughly $(\pi n)^{-1/2}$ for large n.

See `https://github.com/dingyiming0427/CS70-demo/` for a code demo.