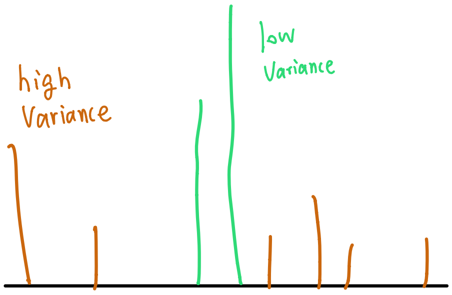
Variance "average squared deviation from mean"



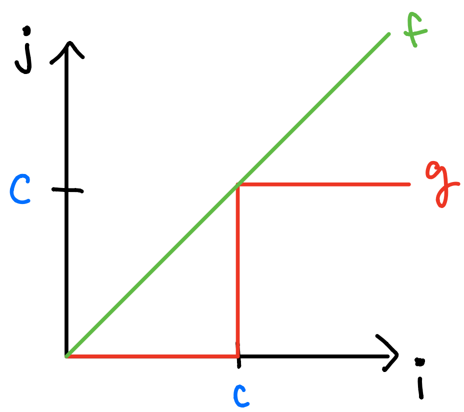
$$E[(X-E[X])^2] = E[X^2 - 2XE[X] + (E[X])^2]$$
$$= E[X^2] - 2E[X]E[X] + (E[X])^2$$
$$= E[X^2] - (E[X])^2$$

Concentration Inequalities

For indep. X, Y $\quad Var(aX+bY) = a^2 Var(X) + b^2 Var(Y)$

Markov (non-negative X)

$$P(X \geq c) \leq \frac{E[X]}{c}$$



$g(i) \leq f(i)$ everywhere

$$\underbrace{\sum_{i=0}^{\infty} g(i)\,P(X=i)}_{\sum_{i=0}^{c} 0\cdot P(X=i) + \sum_{i=c}^{\infty} c\,P(X=i)} \leq \underbrace{\sum_{i=0}^{\infty} f(i)\,P(X=i)}_{\sum_{i=0}^{\infty} i\,P(X=i)}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$c\,P(X \geq c) \leq E[X]$$

Chebyshev

$$P\big[|X-E[X]| \geq c\big] \leq \frac{Var(X)}{c^2}$$

For any R.V. X, define non-negative R.V. $Y = (X-E[X])^2$

Markov on Y: $\quad P\big(\underbrace{(X-E[X])^2 \geq c^2}_{|X-E[X]| \geq c}\big) \leq \dfrac{E[(X-E[X])^2]}{c^2}$

$$P(|X-E[X]| \geq c) \leq \frac{Var(X)}{c^2}$$

Weak Law of Large Numbers (i.i.d. X_i's with $E[X_i] = \mu$)

$$P\left[\left|\frac{1}{n}(X_1+X_2+\dots+X_n) - \mu\right| < \varepsilon\right] \xrightarrow[\text{as } n\to\infty]{} 1$$

Define $Z = \frac{1}{n}(X_1+X_2+\dots+X_n)$

$E[Z] = \mu$

Linearity of Exp.

$$Var(Z) = \frac{1}{n^2}\sum_{i=1}^{n} Var(Z_i) = \frac{Var(Z_i)}{n}$$

Independence

Chebyshev on Z:

$$P\left[\left|\frac{1}{n}(X_1+X_2+\dots+X_n) - \mu\right| \geq \varepsilon\right] \leq \frac{Var(Z_i)}{n\varepsilon^2} \xrightarrow[\text{as } n\to\infty]{} 0$$

1 Variance

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

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2$$

$$= \frac{1}{6}(1^2+2^2+3^2+\dots+6^2) - \left(\frac{1}{6}\cdot(1+2+3+\dots+6)\right)^2$$

$$\frac{91}{6} - \left(\frac{7}{2}\right)^2 = \boxed{\frac{35}{12}}$$

(b) Let Z be a random variable representing the average of n $\underbrace{\text{rolls of a fair die 6-sided}}_{X_1, X_2, \dots, X_n}$ die. What is $\mathrm{Var}(Z)$?

$\underbrace{X_1, X_2, \dots, X_n}_{\text{iid}}$

$$Z = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\mathrm{Var}(Z) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\frac{35}{12} = \frac{1}{n^2}\frac{35}{12}\cdot n = \boxed{\frac{35}{12n}}$$

(c) A building has n floors numbered $1,2,\dots,n$, plus a ground floor G. At the ground floor, m

$$(X_1+\dots+X_n)(X_1+\dots+X_n) = \sum_{i=1}^{n} X_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{n} X_iX_j$$

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.)

$X_1, \dots, X_n$ $\qquad X_i \sim I\{\text{event that elevator doesn't stop at floor } i\}$

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

$$E[Z^2] = E\left(\left[\sum_{i=1}^{n} X_i\right]^2\right) = \sum_{i=1}^{n} E[X_i^2] + \sum_{\substack{i,j=1\\ i\neq j}}^{n} E[X_iX_j]$$

$$n\left(1-\frac{1}{n}\right)^m + n(n-1)\left(1-\frac{2}{n}\right)^m$$

2 Inequality Practice

$$\mathrm{Var}(Z) = E[Z^2] - (E[Z])^2$$

(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.

$X \geq -1$

$Y \geq 4$

$Y = X+5 \qquad Y > 0$

Apply Markov on Y:

$$P(Y \geq 4) \leq \frac{E[Y]}{4}$$

$$= \frac{5+E[X]}{4} = \boxed{\frac{1}{2}}$$

(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.

$$Z = -Y + 10$$

Markov on Z

(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 $\text{Var}(Z)$. Then use Chebyshev's inequality to bound the probability of the sum Z being greater than 400 or less than 300.

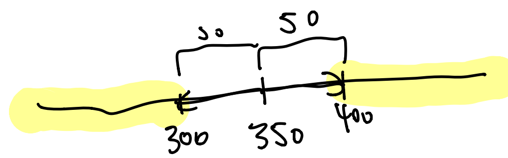
$X_1, \ldots, X_{100}$

$$Z = \sum_{i=1}^{n} X_i$$

$$E[Z] = \sum_{i=1}^{n} E[X_i] = 350$$

$$Var(Z) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} \frac{35}{12} = \frac{35}{12} \cdot n = \frac{3500}{12}$$

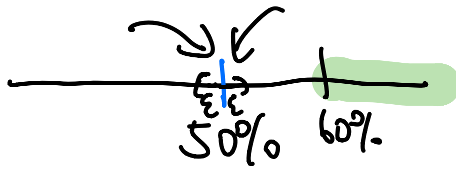
$$P(Z \geq 400 \text{ or } Z \leq 300)$$



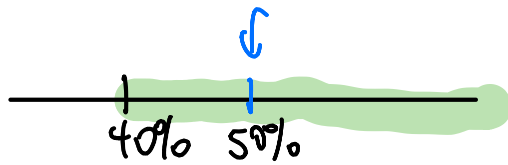
$$\underline{P(|Z - E[Z]| \geq 50)} \leq \frac{Var(Z)}{50^2} = \frac{3500}{12 \cdot 50^2} = \boxed{\frac{7}{60}}$$

3 Working with the Law of Large Numbers

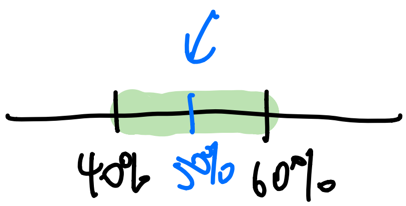
(a) A fair coin is tossed multiple times and you win a prize if there are more than 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.



(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.

$$\binom{10}{5}\frac{1}{2^{10}} \qquad \binom{100}{50}\frac{1}{2^{100}}$$

$$p_n : \binom{2n}{n} \cdot \frac{1}{2^{2n}}$$

$$p_{n+1} : \binom{2n+2}{n+1} \cdot \frac{1}{2^{2n+2}}$$

$$\frac{P_{n+1}}{P_n} = \frac{\binom{2n+2}{n+1} \frac{1}{\cancel{2^{2n+2}}}}{\binom{2n}{n} \frac{1}{\cancel{2^{2n}}}}$$

$$= \frac{\dfrac{(2n+2)(2n+1)\cancel{(2n)!}}{(n+1)\cancel{n!} \cdot (n+1)\cancel{n!}} \cdot \dfrac{1}{4}}{\dfrac{\cancel{(2n)!}}{\cancel{n!} \cdot \cancel{n!}}}$$

$$= \frac{\cancel{2}\cancel{(n+1)}(2n+1)}{(n+1)^{\cancel{2}}} \cdot \frac{1}{\cancel{4}\, 2}$$

$$= \frac{2n+1}{2n+2} < 1$$