$Var(X) = E(X^2) - E(X)^2$

CS 70	Discrete Mathematics and Probability Theory	DIS 5D
Summer 2018	Sinho Chewi and Vrettos Moulos	

1 Variance

This problem will give you practice using the "standard method" to compute the variance of a sum of random variables that are not pairwise independent. Recall that $\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

(a) A building has n floors numbered $1,2,\ldots,n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each person gets off at one of the n floors uniformly at random (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, but the former is a little easier to compute.)

$X =$ # floors elevator does not stop

$= X_1 + X_2 + X_3 + \cdots X_n$

$X_i = \begin{cases} 1 & \text{if nobody gets off} \\ 0 & \text{else} \end{cases}$

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

$E(X) = n \cdot \left(1 - \frac{1}{n}\right)^m$ since $E(X) = E(X_1 + X_2 \cdots + X_n)$

$(E(X))^2 = \left(n \cdot \left(1 - \frac{1}{n}\right)^m\right)^2$

$Var(X) = Var(X_1 + \cdots X_n) = E(X^2) - E(X)^2$

$E(X^2) = E(X_1 + X_2 + \cdots + X_n) = \underbrace{\sum_{i}^{n} E(X_i^2)}_{n \cdot \left(\frac{n-1}{n}\right)^m} + \underbrace{\sum_{i \neq j} E(X_i X_j)}_{n(n-1) \text{ of these}}$

$X_i X_j$: 1 if X_i and $X_j = 1$, 0 else

$\underbrace{E(X_i^2)}_{n \text{ of these}} = 1^2 \cdot P(X_i = 1) = \left(\frac{n-1}{n}\right)^m \rightarrow n \cdot \left(\frac{n-1}{n}\right)^m$

$P(X_i X_j = 1) = \left(\frac{n-2}{n}\right)^m$

(b) A group of three friends has n books they would all like to read. Each friend (independently of the other two) picks a random permutation of the books and reads them in that order, one book per week (for n consecutive weeks). Let X be the number of weeks in which all three friends are reading the same book. Compute $\text{var}(X)$.

2 Will I Get My Package?

A delivery guy in some company is out delivering n packages to n customers, where $n \in \mathbb{N}$, $n > 1$. Not only does he hand a random package to each customer, he opens the package before delivering it with probability $1/2$. Let X be the number of customers who receive their own packages unopened.

(a) Compute the expectation $\mathbb{E}(X)$.

$$X = X_1 + X_2 + \cdots + X_n \Rightarrow E(X) = E(X_1 + X_2 + \cdots + X_n)$$

$$= \underbrace{E(X_1)}_{\frac{1}{2n}} + \underbrace{E(X_2)}_{\frac{1}{2n}} + \cdots + \underbrace{E(X_n)}_{\frac{1}{2n}}$$

$$X_i = \begin{cases} 1 & \text{if gets their package + unopened} \\ 0 & \text{else} \end{cases}$$

$$E(X_i) = P(X_i = 1) = \frac{1}{2n} \qquad \boxed{E(X) = n\frac{1}{2n} = \frac{1}{2}}$$

(b) Compute the variance $\mathrm{var}(X)$. $= E(X^2) - \underbrace{E(X)^2}_{\frac{1}{4}}$

$$E(X^2) = \underbrace{\sum_i^n E(X_i^2)}_{} + \sum_{i \neq j} E(X_i X_j)$$

$$n\left(\frac{1}{2n}\right) + n(n-1)\left(\frac{1}{2n} \cdot \frac{1}{2(n-1)}\right) = \frac{3}{4}$$

(same as $E(X_i)$)

$$Var(X) = E(X^2) + E(X)^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

1 if $X_i = X_j = 1$

two people both got unopened + their package

$$P(X_i X_j = 1) = P(X_i = 1) \cdot P(X_j = 1 \mid X_i = 1)$$

$$= \frac{1}{2n} \cdot \frac{1}{2(n-1)}$$

3 Binomial Conditioning

Let $n \in \mathbb{Z}_+$ and $p, q \in [0,1]$. Let $\boxed{X \sim \text{Binomial}(n,p)}$ and suppose that conditioned on $X = x$, $\boxed{Y \sim \text{Binomial}(x,q)}$. What is the unconditional distribution of Y?

$$P(Y=y) = \sum_{x=0}^{n} P(X=x, Y=y) = \sum_{x=0}^{n} P(X=x) \cdot P(Y=y \mid X=x)$$

$$= \sum_{x=y}^{n} P(X=x)\, P(Y=y \mid X=x) = \sum_{x=y} \binom{n}{x} p^x (1-p)^{n-x} \binom{x}{y} q^y (1-q)^{x-y}$$

$$= \sum_{x=y}^{n} \frac{n!}{x!(n-x)!} \frac{x!}{y!(x-y)!} p^x (1-p)^{n-x} q^y (1-q)^{x-y}$$

$$= \frac{n!}{y!(n-y)!} (pq)^y \sum_{x=y}^{n} \frac{(n-y)!}{(n-x)!(x-y)!} (p(1-q))^{x-y} (1-p)^{n-x}$$

$$= \binom{n}{y} (pq)^y \sum_{x=0}^{n-y} \frac{(n-y)!}{(n-y-x)!\,x!} (p(1-q))^x (1-p)^{n-y-x}$$

$$= \binom{n}{y} (pq)^y \underbrace{\sum_{x=0}^{n-y} \binom{n-y}{x} (p(1-q))^x (1-p)^{n-y-x}}_{(p(1-q)+(1-p))^{n-y}} = \binom{n}{y} (pq)^y (1-pq)^{n-y}$$