Stat 134: Indicator and Covariance Review

Adam Lucas

December 5th, 2018

Conceptual Review

a. What is the computational formula for Var$(X + Y)$?

b. Suppose X is the sum of n identical indicators I_j's. What is Var(X)?

a. $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X,Y)$

$\text{Var}(aX+bY) = a^2\,\text{Var}(X) + b^2\,\text{Var}(Y) + 2ab\,\text{Cov}(X,Y)$

$a, b \in \mathbb{R}$

b. $X = \sum_{j=1}^{n} I_j \qquad \text{Var}(X) = E(X^2) - [E(X)]^2$

$= nE(I_1) + n(n-1)E(I_1 I_2) - [nE(I_1)]^2$

Problem 1

In a bin, there are r red balls and b blue balls. Suppose I take the balls out, one by one (i.e. without replacement), until there are no more red balls in the bin. Let X denote the number of balls taken out. Find:

$X = r \text{ Red} + \underbrace{\#\text{ of Blue before last Red}}_{Y}$

a. $\mathbb{E}(X)$;

b. Var(X).

a. $E(X) = r + E(Y)$

Let I_j be indicator that jth Blue comes before last Red

$E(Y) = E\left(\sum_{j=1}^{b} I_j\right) = \sum_{j=1}^{b} E(I_j) = b \cdot E(I_1) = b\frac{r}{r+1}$

b. $\text{Var}(X) = \text{Var}(r+Y) = \text{Var}(Y) = E(Y^2) - [E(Y)]^2$

$= E(Y) + b(b-1)E(I_1 I_2) - [E(Y)]^2$

$E(I_1 I_2) = P(I_1 = 1, I_2 = 1) = P(I_1 = 1)\,P(I_2 = 1 \mid I_1 = 1) = \frac{r}{r+1}\frac{r+1}{r+2} = \frac{r}{r+2}$

$\text{Var}(X) = \frac{br}{r+1} + b(b-1)\frac{r}{r+2} + \left(\frac{br}{r+1}\right)^2$

Problem 2

Toss a p-coin n times. Let W_r refer to the number of trials until the r_{th} head. Find $Corr(W_1, W_r)$.

$W_1 \sim Geom(p)$ $\quad W_r = W_1 + (W_r - W_1) \sim NegBin(r, p)$

$W_r - W_1 \perp\!\!\!\perp W_1$ by indep. trials

(Recall NegBin is sum of r i.i.d. Geom(p)'s)

$Cov(W_1, W_r) = Cov(W_1, W_1 + (W_r - W_1)) = Cov(W_1, W_1) + Cov(W_1, W_r - W_1)$

$= Var(W_1) = \frac{1-p}{p^2}$

$Corr(W_1, W_r) = \frac{Cov(W_1, W_r)}{SD(W_1)SD(W_r)} = \frac{\frac{1-p}{p^2}}{\sqrt{\frac{1-p}{p^2}}\sqrt{\frac{r(1-p)}{p^2}}} = \frac{1}{\sqrt{r}}$

Problem 3

A p-coin is a coin that lands heads with probability p. Flip a p-coin n times. A "run" is a maximal sequence of consecutive flips that are all the same. For example, the sequence $HTHHHTTH$ with $n = 8$ has five runs, namely H, T, HHH, TT, H. Let X denote the number of runs in these n flips. Find $\mathbb{E}(X)$.

Let I_j be the indicator that jth & $j-1$th trials are different

$X = 1 + \sum_{j=2}^{n} I_j$ (X increments at start of new run)
(1st trial always a start of new run)

$E(X) = 1 + \sum_{j=2}^{n} E(I_j)$

$E(I_j) = pq + qp = 2pq \Rightarrow P(2$ trials w/ diff outcomes, i.e. HT & TH)

$E(X) = 1 + (n-1)2pq$