Statistics 430
HW #11 Solutions

Emil Pitkin

December 8, 2010

1 General principle

This homework relies heavily on indicator functions. To remove clutter, I'll present here a simple calculation that recurs often, but just here, once, and not again.

Whenever we define indicator functions ($X_i = 1$ if some condition is met, 0 otherwise), we are usually interested in $E[\sum_{i=1}^{n} X_i]$. This equals $\sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} P(X_i)$. After defining the indicator functions, we will proceed directly to calculating $P(X_i)$.

2 Ch 7 Problem 21

a

For each of the 365 days in the year, define

$$X_i = \begin{cases} 1 & \text{3 people have birthdays on day } i \\ 0 & \text{otherwise} \end{cases}$$

The number of birthdays on a given day is a binomial random variable, with $n = 100$ and $p = \frac{1}{365}$.

$$P(X_i = 1) = \binom{100}{3}\left(\frac{1}{365}\right)^3\left(\frac{364}{365}\right)^{96}$$

$$365 * \binom{100}{3}\left(\frac{1}{365}\right)^3\left(\frac{364}{365}\right)^{96} = \boxed{.9301}$$

b

For each of the 365 days in the year, define

$$X_i = \begin{cases} 1 & \text{at least one person has a birthday on day } i \\ 0 & \text{otherwise} \end{cases}$$

The probability that at least one birthday falls on day i is 1 minus the probability that no birthdays fall on day i, i.e., 1 minus the probability that all 100 birthdays fall on the other 364 days. If you like, the number of birthdays on a given day is a binomial random variable, with $n = 100$ and $p = \frac{1}{365}$, and we want 1 - P(X=0).

$$\begin{aligned} P(X_i = 1) &= 1 - P(X_i = 0) \\ &= 1 - \binom{100}{0}\left(\frac{1}{365}\right)^0\left(\frac{364}{365}\right)^{100} \\ 365 * \left(1 - \left(\frac{364}{365}\right)^{100}\right) &= \boxed{87.58} \end{aligned}$$

3 Ch 7 Problem 22

This is a thinly veiled version of example $2i$, the coupon collecting problem. We need two facts:

- After we've rolled i distinct numbers, the number of rolls needed to roll a new number is a Geo($\frac{6-i}{6}$) random variable. The probability of a "success" - rolling a new number - is $\frac{6-i}{6}$, since any of the $6-i$ previously unseen numbers, if rolled, will qualify as a new number. The search for the next new number terminates as soon as we see one. Hence the distribution is geometric with the aforementioned parameter.
- The expectation of a Geo(p) random variable is $\frac{1}{p}$

Answer: $\frac{6}{6} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = \boxed{14.7}$

4 Ch 7 Problem 23

Take the hint. With X_i and Y_i defined as in the hint, the number of white balls seen is $\sum_{i=1}^{5} X_i + \sum_{i=1}^{8} Y_i$. When it comes time to withdraw the balls, we have 20 balls to withdraw (the original 18 + the 2 which migrated). We withdraw 3 balls, so $P(Y_i = 1) = \frac{3}{20}$. $P(X_i = 1) = P(X_i = 1|\text{ball i is transferred})P(\text{ball i is transferred}) + P(X_i = 1|\text{ball i is not transferred})P(\text{ball i is not transferred}) = \frac{2}{11} * \frac{3}{20} + 0 = \frac{3}{110}$. Hence the expected number of white balls withdrawn is

$$\begin{aligned} 5 * P(X_i = 1) + 8 * P(Y_i = 1) &= 5 * \frac{3}{110} + 8 * \frac{3}{20} \\ &= \boxed{\tfrac{147}{110} = 1.336} \end{aligned}$$

5 Ch 7 Problem 35

In each of the following problems, let X denote the number of cards that need to be flipped over in order for r cards of a particular type to have been seen. X follows a negative hypergeometric distribution (example $3e$), with parameters n (= number of special cards, i.e., spades) and m (= $52 - n$). We know that $E[X] = \frac{r(n+m+1)}{n+1}$. Since $n + m = 52$, $E[X] = \frac{r*53}{n+1}$

a

Here, $r = 2, n = 4$. So $E[X] = \frac{2*(53)}{5} = \boxed{21.2}$

b

Here, $r = 5, n = 13$. So $E[X] = \frac{5*(53)}{14} = \boxed{18.929}$

c

Here, $r = 13, n = 13$. So $E[X] = \frac{13*(53)}{14} = \boxed{49.214}$

6 Ch 7 Problem 36

Let

$$X_i = \begin{cases} 1 & \text{the ith roll is a 1} \\ 0 & \text{otherwise} \end{cases}$$

$$Y_i = \begin{cases} 1 & \text{the ith roll is a 2} \\ 0 & \text{otherwise} \end{cases}$$

Notice that X, the number of 1's, is equal to $\sum_{i=1}^{n} X_i$, with Y represented similarly. We are interested in $Cov[X, Y] = E[XY] - E[X]E[Y]$. $E[X] = E[Y] = \frac{n}{6}$,

$$\begin{aligned} E[XY] &= E[\sum_{i=1}^{n} X_i * \sum_{i=1}^{n} Y_i] \\ &= E[\sum_{i=1}^{n} X_i Y_i] + E[\sum_{i=1}^{n} \sum_{i \neq j} X_i Y_j] \\ &= 0 + \frac{n(n-1)}{36} = \frac{n^2}{36} - \frac{n}{36} \end{aligned}$$

So

$$\begin{aligned} E[XY] - E[X]E[Y] &= \frac{n^2}{36} - \frac{n}{36} - \frac{n}{36}\frac{n}{36} \\ &= \boxed{-\frac{n}{36}} \end{aligned}$$

7 Ch 7 Problem 37

Define F to be the outcome of the first roll, and S to be the outcome of the second roll. Since X=F+S and Y=F-S, we are asked for $Cov[F + S, F - S]$. Keep in mind that $E[F - S] = E[F] - E[S] = 3.5 - 3.5 = 0$, and that $E[F^2 - S^2] = E[F^2] - E[S^2] = 0$, since the rolls are identically distributed.

$$\begin{aligned} Cov[F + S, F - S] &= E[(F + S)(F - S)] - E[F + S]E[F - S] \\ &= E[F^2 - S^2] - E[F + S]E[F - S] \\ &= 0 - 0 = \boxed{0} \end{aligned}$$

This is a classic problem which I had when I took probability. The answer would be the same if we were asked for the covariance between the sum and difference of the two rolls.

8 Ch 7 Problem 42

a

$$
\begin{aligned}
X_i &= \begin{cases} 1 & \text{the ith pair consists of a man and a woman} \\ 0 & \text{otherwise} \end{cases} \\
P(X_i = 1) &= \frac{\binom{10}{1}\binom{10}{1}}{\binom{20}{2}} \\
10 * \frac{\binom{10}{1}\binom{10}{1}}{\binom{20}{2}} &= \boxed{\tfrac{100}{19} = 5.26}
\end{aligned}
$$

Variance time.

$$
\begin{aligned}
Var[X] &= Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i] + \sum_{i=1}^{n} \sum_{i \neq j} Cov[X_i, X_j] \\
Var[X_i] &= E[X_i](1 - E[X_i]) = \frac{10}{19} * \frac{9}{19} \\
Cov[X_i, X_j] &= E[X_i X_j] - E[X_i]E[X_j] = E[X_i X_j] - \left(\frac{10}{19}\right)^2 \\
E[X_i X_j] &= P(X_i = 1, X_j = 1) = \frac{\binom{10}{1}\binom{10}{1}\binom{9}{1}\binom{9}{1}}{\binom{20}{2}\binom{18}{2}} = \frac{90}{19 * 17} \\
Var[X] &= 10 * \left[\frac{10}{19}\frac{9}{19}\right] + 90\left[\frac{90}{19 * 17} - \frac{100}{361}\right] \\
&= \boxed{2.6397}
\end{aligned}
$$

b

$$
\begin{aligned}
X_i &= \begin{cases} 1 & \text{the ith pair consists of a man and his wife} \\ 0 & \text{otherwise} \end{cases} \\
P(X_i = 1) &= \frac{10}{\binom{20}{2}} \\
10 * \frac{10}{\binom{20}{2}} &= \boxed{\tfrac{10}{19} = .526}
\end{aligned}
$$

Round two of variance time.

$$\begin{aligned}
Var[X] &= Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i] + \sum_{i=1}^{n}\sum_{i\neq j} Cov[X_i, X_j] \\
Var[X_i] &= E[X_i](1 - E[X_i]) = \frac{1}{19}\frac{18}{19} \\
Cov[X_i, X_j] &= E[X_iX_j] - E[X_i]E[X_j] = E[X_iX_j] - \left(\frac{1}{19}\right)^2 \\
E[X_iX_j] &= P(X_i = 1, X_j = 1) = \frac{\binom{10}{1}\binom{9}{1}}{\binom{20}{2}\binom{18}{2}} = \frac{90}{19 * 17} \\
Var[X] &= 10 * \left[\frac{1}{19}\frac{18}{19}\right] + 90\left[\frac{1}{19 * 17} - \frac{1}{361}\right] \\
&= \boxed{.5279}
\end{aligned}$$

9 Ch 7 Problem 48

a

Since $X \sim Geo(\frac{1}{6})$, $E[X] = \boxed{6}$.

b

If $Y = 1$, we know that the first roll was a 5. Henceforth, the number of rolls until we roll a 6 is a Geo($\frac{1}{6}$) random variable, with expectation 6. So $E[X|Y = 1] = 1 + 6 = \boxed{7}$.

c

Since Y, the first instance of a 5, is equal to 5 in this problem, we know that we could only have rolled 1,2,3,4 or 6 in the first 4 rolls. If the first 6 appears in the first 4 rolls, X is a Geo($\frac{1}{5}$) random variable. If the first 6 appears after the 5th roll, then X is a Geo($\frac{1}{6}$) random variable, with expectation 6.

$$\begin{aligned}
E[X|Y = 5] &= 1 * P(X = 1|Y = 5) + 2 * P(X = 2|Y = 5) + 3 * P(X = 3|Y = 5) \\
&+ 4 * P(X = 4|Y = 5) + (5 + 6)P(X > 5|Y = 5) \\
&= 1 * (.2 * .8^0) + 2 * (.2 * .8^1) + 3 * (.2 * .8^2) \\
&+ 4 * (.2 * .8^3) + (5 + 6)(1 - .2 * .8^0 - .2 * .8^1 - .2 * .8^2 - .2 * .8^3) \\
&= \boxed{5.8192}
\end{aligned}$$

10 Ch 7 Problem 50

$$\begin{aligned}
f_Y(y) &= \int_0^\infty \frac{e^{-x/y}e^{-y}}{y}\,dx \\
&= e^{-y} \\
f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} \\
&= \frac{\frac{e^{-x/y}e^{-y}}{y}}{e^{-y}} \\
&= \frac{e^{-x/y}}{y} \\
E[X^2|Y=y] &= \int_0^\infty \frac{x^2e^{-x/y}}{y}\,dx \quad \text{integrate by parts} \\
&= \boxed{2y^2}
\end{aligned}$$

11 Ch 7 Problem 58

a

Let F be the number of flips. Condition on the outcome of the first flip. If the first flip is heads, then the number of subsequent flips until tails is a Geo(1-p) random variable, with expectation $\frac{1}{1-p}$. So the expected number of total flips until heads and tails are flipped is $1 + \frac{1}{1-p}$.

$$\begin{aligned}
E[F] &= E[F|H_1]P(H_1) + E[F|T_1]P(T_1) \\
&= \left(1 + \frac{1}{1-p}\right)p + \left(1 + \frac{1}{p}\right)(1-p) \\
&= \boxed{\tfrac{1}{p(1-p)} - 1}
\end{aligned}$$

b

The last flip lands on heads if the first flip lands on tails, which occurs with probability $\boxed{1-p}$.