Homework_13

December 4, 2019

```
In [2]: # HIDDEN
        from datascience import *
        from prob140 import *
        import numpy as np
        import matplotlib.pyplot as plt
        plt.style.use('fivethirtyeight')
        %matplotlib inline
        from scipy import stats
```

1 Homework 13

1.0.1 1. Waiting Till HH

In a lab earlier in the term, you found a formula for the expected waiting time till any finite pattern of heads and tails appears in a sequence of coin tosses. In this exercise you will revisit that calculation in the case of a simple pattern HH, and find the variance of the waiting time as well.

A p-coin is tossed repeatedly. Let W_H be the number of tosses till the first head appears, and W_{HH} the number of tosses till two consecutive heads appear.

(a) Describe a random variable X that depends only on the tosses after W_H and satisfies $W_{HH} = W_H + X$.

(b) Use Part (a) to find $E(W_{HH})$. What is its value when $p = 1/2$?

(c) Use Parts (a) and (b) to find $Var(W_{HH})$. What is the value of $SD(W_{HH})$ when $p = 1/2$?

1.0.2 Solution

(a) $X = 1$ with probability p.

With probability q, $X = 1 + W^*_{HH}$ where W^*_{HH} is an independent copy of W_{HH}.

(b) Let $\mu = E(W_{HH})$ and let I be the indicator of heads on the toss after W_H. Then

$E(X \mid I) = 1$ with probability p and $1 + \mu$ with probability q.
By iteration,
$E(X) = p + q(1 + \mu) = 1 + q\mu$.
Since $W_{HH} = W_H + X$, we have $\mu = E(W_H) + E(X) = \frac{1}{p} + 1 + q\mu$.
So $p\mu = \frac{1}{p} + 1$ and hence $\mu = \frac{1}{p^2} + \frac{1}{p}$.
When $p = 1/2$, $\mu = 4 + 2 = 6$.

(c) For brevity, let $\sigma^2 = Var(W_{HH})$.

$\sigma^2 = Var(W_H) + Var(X)$ since W_H and X are functions of disjoint sets of tosses.
$Var(W_H) = \frac{q}{p^2}$ since W_H is geometric. So $\sigma^2 = \frac{q}{p^2} + Var(X)$.
Find $Var(X)$ by conditioning on I, the indicator of the toss after W_H landing heads.
$Var(X) = E(Var(X \mid I)) + Var(E(X \mid I))$
The first piece:
$Var(X \mid I = 1) = 0$ and $Var(X \mid I = 0) = \sigma^2$.
So $E(Var(X \mid I) = 0 \cdot p\ +\ \sigma^2 q = \sigma^2 q$.
The second piece:
From (b), $E(X \mid I) = 1$ with probability p and $1 + \mu$ with probability q, and $E(E(X \mid I)) = E(X) = 1 - q\mu$.
So $Var(E(X \mid I)) = 1^2 p + (1+\mu)^2 q - (1+q\mu)^2$
$= 1 + 2\mu q + \mu^2 q - 1 - 2q\mu - q^2\mu^2$
$= \mu^2 qp$
So $Var(X)\ =\ \sigma^2 q + \mu^2 qp$
Plug this into $\sigma^2 = \frac{q}{p^2} + Var(X)$ to get
$\sigma^2 = \frac{q}{p^2} + \sigma^2 q + \mu^2 qp$
So $p\sigma^2 = \frac{q}{p^2} + \mu^2 qp$ and so
$\sigma^2 = \frac{q}{p^3} + \mu^2 q$
That's fine as an answer. But it simplifies after substituting $\mu = \frac{1}{p^2} + \frac{1}{p}$:
$\sigma^2 = q\left(\frac{1}{p^3} + \frac{1}{p^4} + \frac{2}{p^3} + \frac{1}{p^2}\right)$
$= (1-p)\left(\frac{3}{p^3} + \frac{1}{p^4} + \frac{1}{p^2}\right)$
$= \frac{1}{p^4} + \frac{2}{p^3} - \frac{2}{p^2} - \frac{1}{p}$
When $p = 1/2$, $\sigma^2 = 16 + 16 - 8 - 2 = 22$, so $\sigma \approx 4.7$.

1.0.3 2. Random Vector Workout

A random vector $\mathbf{Y} = [Y_1\ Y_2\ \cdots\ Y_n]^T$ has mean vector $\boldsymbol{\mu}$ and covariance matrix $\sigma^2\mathbf{I}_n$ where $\sigma > 0$ is a number and $\mathbf{I}_n$ is the $n \times n$ identity matrix.

(a) Pick one option and explain: Y_1 and Y_2 are

(i) independent. (ii) uncorrelated but might not be independent. (iii) not uncorrelated.

(b) Pick one option and explain: $Var(Y_1)$ and $Var(Y_2)$ are

(i) equal. (ii) possibly equal, but might not be. (iii) not equal.

(c) For $m \leq n$ let $\mathbf{A}$ be an $m \times n$ matrix of real numbers, and let $\mathbf{b}$ be an $m \times 1$ vector of real numbers. Let $\mathbf{V} = \mathbf{AY} + \mathbf{b}$. Find the mean vector $\boldsymbol{\mu}_{\mathbf{V}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{V}}$ of $\mathbf{V}$.

(d) Let $\mathbf{c}$ be an $m \times 1$ vector of real numbers and let $W = \mathbf{c}^T\mathbf{V}$ for $\mathbf{V}$ defined in Part (c). In terms of $\mathbf{c}$, $\boldsymbol{\mu}_{\mathbf{V}}$ and $\boldsymbol{\Sigma}_{\mathbf{V}}$, find $E(W)$ and $Var(W)$.

(a) (ii) because the (1,2) element of the covariance matrix is 0

(b) (i) because all the diagonal elements of the covariance matrix are σ^2

(c) $\boldsymbol{\mu}_{\mathbf{V}} = \mathbf{A}\boldsymbol{\mu} + \mathbf{b}$, $\boldsymbol{\Sigma}_{\mathbf{V}} = \mathbf{A}\sigma^2\mathbf{I}n\mathbf{A}^T = \sigma^2\mathbf{A}\mathbf{A}^T$

(d) $E(W) = \mathbf{c}^T\boldsymbol{\mu}_{\mathbf{V}}$, $Var(W) = \mathbf{c}^T\boldsymbol{\Sigma}_{\mathbf{V}}\mathbf{c}$

1.0.4 3. Normals and Coins

Let X be standard normal. Construct a random variable Y as follows:

- Toss a fair coin.
- If the coin lands heads, let $Y = X$.
- If the coin lands tails, let $Y = -X$.

(a) Find the cdf of Y.

(b) Find $E(XY)$ by conditioning on the result of the toss.

(c) Are X and Y uncorrelated?

(d) Are X and Y independent?

(e) Is the joint distribution of X and Y bivariate normal?

(a) For all y, $\Phi(y) = 1 - \Phi(-y)$.

Let I be the indicator of heads.
$P(Y < y) = P(Y < y, I = 1) + P(Y < y, I = 0) = (1/2)P(X < y) + (1/2)P(-X < y) = (1/2)\Phi(y) + (1/2)(1 - \Phi(-y)) = (1/2)\Phi(y) + (1/2)\Phi(y) = \Phi(y)$
So Y is standard normal.

(b) Given $I = 1$, $XY = X^2$ so $E(XY \mid I = 1) = E(X^2) = 1$.

Given $I = 0$, $XY = -X^2$ so $E(XY \mid I = 0) = -E(X^2) = -1$.
So $E(XY) = (1/2)1 + (1/2)(-1) = 0$

(c) Yes by (b).

(d) No. For example if $X = 2$ then Y has to be either 2 or -2.

(e) No. If they were bivariate normal then they would be independent because they are uncorrelated. But they aren't independent.

1.0.5 4. Correlation

The covariance of random variables X and Y has nasty units: the product of the units of X and the units of Y. Dividing the covariance by the two SDs results in an important pure number.

The *correlation coefficient* between random variables X and Y is defined as

$$r(X,Y) \;=\frac{Cov(X,Y)}{SD(X)SD(Y)}$$

It is called the correlation, for short. The definition explains why X and Y are called *uncorrelated* if $Cov(X,Y) = 0$.

a) Let X^* be X in standard units and let Y^* be Y in standard units. Check that

$$r(X,Y) = E(X^*Y^*)$$

This is the random variable version of the Data 8 definition of the correlation between two data variables: convert each variable to standard units; multiply each pair; take the mean of the products.

b) Use the fact that $(X^* + Y^*)^2$ and $(X^* - Y^*)^2$ are non-negative random variables to show that $-1 \le r(X,Y) \le 1$.

[First find the numerical values of $E(X^*)$ and $E(X^{*2})$. Then find $E(X^* + Y^*)^2$.]

c) Show that if $Y = aX + b$ where $a \neq 0$, then $r(X,Y)$ is 1 or -1 depending on whether the sign of a is positive or negative.

d) Consider a sequence of i.i.d. Bernoulli (p) trials. For any positive integer k let X_k be the number of successes in trials 1 through k. **Use bilinearity** to find $Cov(X_n, X_{n+m})$ and hence find $r(X_n, X_{n+m})$.

e) Fix n and find the limit of your answer to **c** as $m \to \infty$. Explain why the limit is consistent with intuition.

1.0.6 Solution

a) $E(X^*) = E(Y*) = 0$

$SD(X^*) = SD(Y*) = 1$

$Cov(X^*, Y^*) = \frac{1}{SD(X)SD(Y)} Cov(X - E(X), Y - E(Y)) = \frac{1}{SD(X)SD(Y)} Cov(X,Y) = r(X,Y)$

$r(X,Y) = Cov(X^*, Y^*) = E(X^*Y^*) - E(X^*)E(Y^*) = E(X^*Y^*) - 0 = E(X^*Y^*)$

b)

$E(X^*) = 0$

$E(X^{*2}) = Var(X^*) + (E(X^*))^2 = 1 + 0 = 1$

$0 \le E[(X^* + Y^*)^2] = E(X^*2) + E(X^*2) + 2E(X^*Y^*) = 1 + 1 + 2E(X^*Y^*) = 2 + 2E(X^*Y^*)$

$-2 \le 2E(X^*Y^*)$

$-1 \le E(X^*Y^*)$

$0 \le E[(X^* - Y^*)^2] = E(X^*2) + E(X^*2) - 2E(X^*Y^*) = 1 + 1 - 2E(X^*Y^*) = 2 - 2E(X^*Y^*)$

$2E(X^*Y^*) \le 2$

$E(X^*Y^*) \le 1$

$-1 \le E(X^*Y^*) = r(X,Y) \le 1$

c) $Cov(X, aX + b) = aCov(X,X) + Cov(b,X) = aVar(X)$

$Var(aX + b) = a^2 Var(X)$

$r(X,Y) = r(X, aX+b) = \frac{aVar(X)}{|a|SD(X)SD(X)} = \frac{a}{|a|}$

The above is 1 is $a > 0$, and -1 if $a < 0$.

d)

$Cov(X_n, X_{n+m}) = Cov(X_n, X_n + X_{n+1,n+m}) = Cov(X_n, X_n) + Cov(X_n, X_{n+1,n+m})$

$= Var(X_n) + 0 = Var(X_n) = np(1-p)$

$r(X_n, X_{n+m}) = \frac{Cov(X_n, X_{n+m})}{SD(X_n)SD(X_{n+m})} = \frac{np(1-p)}{\sqrt{np(1-p)}\sqrt{(n+m)p(1-p)}} = \frac{np(1-p)}{p(1-p)\sqrt{n(n+m)}}$

$= \frac{n}{\sqrt{n(n+m)}} = \sqrt{\frac{n^2}{n(n+m)}} = \sqrt{\frac{n}{n+m}}$

e) As $m \to \infty$, $\sqrt{\frac{n}{n+m}} \to 0$

This makes sense because as $m \to \infty$, the overlap between X_n and X_{n+m} decreases. The overlapping X_n portion is insignificant because X_n and X_m are independent and m is large compared to n.

1.1 Submission Instructions

Many assignments throughout the course will have a written portion and a code portion. Please follow the directions below to properly submit both portions.

1.1.1 Written Portion

- Scan all the pages into a PDF. You can use any scanner or a phone using an application. Please **DO NOT** simply take pictures using your phone.
- Please start a new page for each question. If you have already written multiple questions on the same page, you can crop the image or fold your page over (the old-fashioned way). This helps expedite grading.
- It is your responsibility to check that all the work on all the scanned pages is legible.

1.1.2 Code Portion

- Save your notebook using File > Save and Checkpoint.
- Generate a PDF file using File > Download as > PDF via LaTeX. This might take a few seconds and will automatically download a PDF version of this notebook.
  - If you have issues, please make a follow-up post on the general HW 13 Piazza thread.

1.1.3 Submitting

- Combine the PDFs from the written and code portions into one PDF. Here is a useful tool for doing so.
- Submit the assignment to Homework 13 on Gradescope.
- **Make sure to assign each page of your pdf to the correct question.**
- **It is your responsibility to verify that all of your work shows up in your final PDF submission.**

1.1.4 We will not grade assignments which do not have pages selected for each question.