Homework_13

December 3, 2019

Probability for Data Science

UC Berkeley, Fall 2019

Ani Adhikari and Jim Pitman

[ ]:
```
# 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 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 $m \le n$ 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 $\mathbf{V} = \mathbf{A}\mathbf{Y} + \mathbf{b}$. 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 $Var(W)$.

1.0.3 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?

1.0.4 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.

1.1 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.

hw 13.

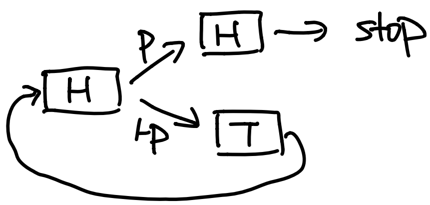
1. → p-coin repeatedly & W_H = # toss till 1^{st} H.

→ W_{HH} = # tosses till HH.

a. → only on tosses after W_H & $W_{HH} = W_H + X$.

⇒ let $Y = W_{HH}$

⇒ after W_H



⇒ let X be a random variable that describe the # of tosses

till pattern head & head (W_{HH}) exist after the 1^{st} head (W_H)

⇒ ① if it's a head

→ $X = 1$

→ probability p

⇒ ② if it's a tail

→ $X = 1 + Y$

→ probability $1-p$

⇒ thus, for random variable X.

probability	value
p	$X=1$
$1-p.$	$X=1+Y$

b. $\rightarrow E(W_{HH})$ & value $p=\frac{1}{2}$

$\Rightarrow W_{HH} = W_H + X$

$\Rightarrow E(W_{HH}) = E(W_H + X)$

$\Rightarrow Y = E(W_H) + E(X)$

$\Rightarrow \quad = \frac{1}{p} + p*1 + (1-p)*(1+Y)$

$= \frac{1}{p} + 1 + qY$

$\Rightarrow pY = \frac{1}{p} + 1$

$\Rightarrow Y = \frac{1}{p^2} + \frac{1}{p}$

$\Rightarrow E(W_{HH}) = \frac{1}{p^2} + \frac{1}{p}$

$\Rightarrow$ for $p = \frac{1}{2}$

$\Rightarrow E(W_{HH}) = \frac{1}{(\frac{1}{2})^2} + \frac{1}{(\frac{1}{2})} = \frac{1}{(\frac{1}{4})} + \frac{1}{(\frac{1}{2})} = 4+2$

$=6$

$\Rightarrow$ thus, $E(W_{HH}) = \frac{1}{p^2} + \frac{1}{p}$, and

at $p=\frac{1}{2}$, $E(W_{HH}) = 6.$

c. $\rightarrow Var(W_{HH})$

$$\Rightarrow X = \begin{cases} 1 & \leftarrow \text{prob } p \ (I_H) \\ Y^* + 1 & \leftarrow \text{prob } 1-p \ (I_T) \end{cases}$$

$$\Rightarrow Var(X) = E(Var(Y|X)) + Var(E(Y|X))$$

$$\Rightarrow E(Var(Y|X)) = 1 * p + q * Var(Y)$$

$$\Rightarrow Var(E(Y|X)) = E(E(Y|X)^2) - [E(E(Y|X))]^2$$

$$= 1^2 * p + (1+E(Y))^2 * q - [E(X)]^2$$

$$= 1^2 * p + \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)^2 - (p + q * (1+E(Y)))^2$$

$$= p + \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)^2 - \left(p + q * \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)\right)^2$$

$$= p + \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)^2 - \left(p + q + \frac{q}{p^2} + \frac{q}{p}\right)^2$$

$$\Rightarrow Var(Y) = Var(W_H) + Var(X)$$

$$\Rightarrow Var(Y) = \frac{q}{p^2} + p + q * Var(Y) + p + \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)^2 - \left(p + q + \frac{q}{p^2} + \frac{q}{p}\right)^2$$

$$\Rightarrow p * Var(Y) = \frac{q}{p^2} + p + p + \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)^2 - \left(p + q + \frac{q}{p^2} + \frac{q}{p}\right)^2$$

$$\Rightarrow Var(Y) = \frac{q}{p^3} + 1 + 1 + p * \left(1 + \frac{1}{p^2} + \frac{1}{p}\right)^2 - p * \left(p + q + \frac{q}{p^2} + \frac{q}{p}\right)^2$$

$$\Rightarrow \text{at } p = \tfrac{1}{2}$$

$$\Rightarrow Var(Y) = \frac{\frac{1}{2}}{\left(\frac{1}{2}\right)^3} + 2 + \frac{1}{2} * \left(1 + \frac{1}{\left(\frac{1}{2}\right)^2} + \frac{1}{\left(\frac{1}{2}\right)}\right)^2 - \frac{1}{2} * \left(1 + \frac{\frac{1}{2}}{\left(\frac{1}{2}\right)^2} + \frac{\frac{1}{2}}{\frac{1}{2}}\right)^2$$

$$= 4 + 2 + \frac{1}{2} * (1 + 4 + 2)^2 - \frac{1}{2} * (1 + 2 + 1)^2$$

$$= 4+2+\frac{1}{2}*49-\frac{1}{2}*16$$

$$= 6+\frac{49}{2}-8 = \frac{49}{2}-2 = \frac{49}{2}-\frac{4}{2}$$

$$= \frac{45}{2}$$

$\Rightarrow Var(Y) = \frac{45}{2}$

$\Rightarrow SD(Y) = \sqrt{\frac{45}{2}} = \frac{3\sqrt{10}}{2}$

$\Rightarrow$ thus, $Var(W_{HH}) = \frac{q}{p^3} + 1 + 1 + p*(1+\frac{1}{p^2}+\frac{1}{p})^2 - p*(p+q+\frac{q}{p^2}+\frac{q}{p})^2$

and at $p=\frac{1}{2}$, $SD(W_{HH}) = \frac{3\sqrt{10}}{2}$.

2. $\rightarrow$ vector $Y = [Y_1 \ Y_2 \ \cdots\cdots \ Y_n]^T$

$\rightarrow$ mean vector μ & cov matrix $\sigma^2 I_n$

a. $\rightarrow$ Y_1 & Y_2

$$\Rightarrow \Sigma_Y = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & & & \\ & 0 & \cdots & \sigma^2 \end{bmatrix} \rightarrow n*n \text{ matrix}$$

$\Rightarrow$ so if $\Sigma(a,b)$ only $\neq 0$ at $a==b$

$\Rightarrow cov(Y_1, Y_2) = \Sigma_{1,2} = 0$

$\Rightarrow$ Y_1 & Y_2 uncorrelated

$\Rightarrow$ we don't know if Y_1 & Y_2 are multivariate normal

$\Rightarrow$ thus, we pick (ii)

that Y_1 & Y_2 are uncorrelated but might not be independent.

b.

$\Rightarrow Var(Y_1) = \Sigma_{(1,1)} = \sigma^2$

$\Rightarrow Var(Y_2) = \Sigma_{(2,2)} = \sigma^2$

$\Rightarrow Var(Y_1) == Var(Y_2)$

$\Rightarrow$ thus, we pick (i)

that $Var(Y_1)$ and $Var(Y_2)$ are equal.

c. $\rightarrow \mathbb{A} = m * n$ matrix & $m < n$

$\rightarrow \mathbb{b} = m * 1$ vector & $\mathbb{V} = \mathbb{A}\mathbb{Y} + \mathbb{b}$

$\rightarrow$ mean vector $\mu_{\mathbb{V}}$ & cov $\Sigma_{\mathbb{V}}$

$\Rightarrow \mu_{\mathbb{Y}} = E[\mathbb{Y}]$

$\Rightarrow \mathbb{V} = \mathbb{A}\mathbb{Y} + \mathbb{b}$

$\Rightarrow \mathbb{V}$ = linear transformation of $\mathbb{Y}$ element.

$\Rightarrow E[\mathbb{b}_i] = \mathbb{b}_i$

$\Rightarrow E(\mathbb{V}_i) = \mathbb{A}_{i*} \mu_{\mathbb{Y}} + \mathbb{b}(i)$

$\Rightarrow \mu_{\mathbb{V}} = \mathbb{A} * \mu_{\mathbb{Y}} + \mathbb{b}$

← chapter 23.

$\Rightarrow \Sigma_V = A \Sigma_Y A^T$

$\Rightarrow$ thus, mean vector $\mu_V = A * \mu_Y + b$, and

covariance matrix $\Sigma_V = A \Sigma_Y A^T$

d. $\longrightarrow C = m*1$ constant vector

$\rightarrow W = C^T * V$ & V in part c.

$\rightarrow E(W)$ & $Var(W)$, C, μ_V, Σ_V

$\Rightarrow C^T = \overset{m}{[\quad\quad]}1$ & $V = \overset{1}{\begin{bmatrix} \\ \\ \end{bmatrix}} n$

$\Rightarrow E(W) = E(C^T * V)$

$= E(C^T) * E(V)$

$= C^T * E(V)$ ← C^T are all constant

$= C^T * \mu_V$

$\Rightarrow W$ is a linear transformation of V elements

$\Rightarrow var(W) = C \Sigma_V C^T$ ← chapter 23

$\Rightarrow$ thus, $E(W) = C^T * \mu_V$,

and $var(W) = C \Sigma_V C^T$

3.

a. $\rightarrow$ cdf. of Y

$\Rightarrow$ cdf of $Y = P(Y \le y)$

$= \frac{1}{2} * P(X \le y) + \frac{1}{2} * P(-X \le y)$

$= \frac{1}{2} * P(X \le y) + \frac{1}{2} * P(X \geq -y)$

$= \frac{1}{2} * \phi(y) + \frac{1}{2} * (1 - \phi(-y))$. $\leftarrow$ standard norm X.

$= \frac{1}{2} * \phi(y) + \frac{1}{2} \phi(y)$ $\leftarrow$ symmetry

$= \phi(y)$ $\leftarrow$ y is also standard normal

$\Rightarrow$ thus, the cdf of Y,

$P(Y \le y) = \phi(y)$

b.

$\Rightarrow E(XY) = E(E(XY \mid \text{tosses}))$

$= \frac{1}{2} E(XY \mid \text{head}) + \frac{1}{2} E(XY \mid \text{tail})$.

$= \frac{1}{2} E(X^2) + \frac{1}{2} E(-X^2)$.

$\Rightarrow$ since X is standard normal dist, $X^2 = \text{gamma}(\frac{1}{2}, \frac{1}{2})$ dist

$\Rightarrow E(X^2) = \frac{1}{2} / \frac{1}{2}$

$= 1$

$\Rightarrow E(-X^2) = -1 * E(X^2) = 1 * -1$

$= -1$

$\Rightarrow E(XY) = \frac{1}{2} E(X^2) + \frac{1}{2} E(-X^2).$

$= \frac{1}{2} * 1 + \frac{1}{2} * -1$

$= 0.$

$\Rightarrow$ thus, $E(XY) = 0.$

c. $\Rightarrow$ uncorrelated

$\Rightarrow$ if X & Y uncorrelated, $\text{cov}(X, Y) = 0$

$\Rightarrow Var(X+Y) = Var(X) + Var(Y) + 2\text{cov}(X, Y)$

$\Rightarrow$ since X & Y are both standard normal dist

$\Rightarrow Var(X+Y) = Var(X) + Var(Y) = 1+1$

$= 2$

$\Rightarrow Var(X) = 1$ & $Var(Y) = 1$

$\Rightarrow 2\text{cov}(X, Y) = Var(X) + Var(Y) - Var(X+Y)$

$= 1+1-2$

$\Rightarrow \text{cov}(X, Y) = 0$

$\Rightarrow$ thus, X and Y are uncorrelated.

d. → independence

⇒ if X & Y are independence, $P(X,Y) = P(X) * P(Y)$ and.

$P(X|Y) = P(X)$ & $P(Y|X) = P(Y)$.

⇒ let $X = x$, they Y can only be x & $-x$

⇒ $P(Y \neq x \text{ or } -x) = 0$ & $P(Y|X) \neq P(Y)$

⇒ thus, X and Y are dependence.

e.

⇒ both X & Y are bivariate normal

⇒ however X & Y are dependent ← part d.

⇒ for joint dist of X & Y to be bivariate normal, X & Y must be linear combination of independent iid. standard normal.

⇒ thus, since X and Y are dependent,

joint dist of X and Y are not bivariate normal.

4. → correlation coefficient $r(X,Y) = \frac{cov(X,Y)}{SD(X)SD(Y)}$

a. → X^* & Y^* = standard unit X, Y.

→ $r(X,Y) = E(X^* Y^*)$

$\rightarrow$ each variable to standard unit & multiple each pair

$\rightarrow$ mean of the product

$\Rightarrow$ let μ_X & μ_Y = mean of X, Y variable

$$\Rightarrow r(X,Y) = \frac{cov(X,Y)}{SD(X)SD(Y)}$$

$$= \frac{E((X-\mu_X)(Y-\mu_Y))}{SD(X)SD(Y)} \leftarrow \text{chapter 13}$$

$\Rightarrow$ since $SD(X)$ & $SD(Y)$ are constants

$$\Rightarrow r(X,Y) = E\left(\frac{(X-\mu_X)}{SD(X)} * \frac{(Y-\mu_X)}{SD(Y)}\right)$$

$$\Rightarrow X^* = \frac{X-\mu_X}{SD(X)} \quad \& \quad Y^* = \frac{Y-\mu_Y}{SD(Y)}$$

$$\Rightarrow E(X^*Y^*) = E\left(\frac{(X-\mu_X)}{SD(X)} * \frac{(Y-\mu_X)}{SD(Y)}\right)$$

$$\Rightarrow r(X,Y) = E(X^*Y^*) = E\left(\frac{(X-\mu_X)}{SD(X)} * \frac{(Y-\mu_X)}{SD(Y)}\right)$$

$\Rightarrow$ thus, $r(X,Y) = E(X^*Y^*)$.

b. $\rightarrow (X^*+Y^*)^2$ & $(X^*-Y^*) > 0$

$\rightarrow -1 \leq r(X,Y) \leq 1$

$\rightarrow$ numerical value of $E(X^*)$ & $E(Y^*)$

$\rightarrow E(X^*+Y^*)^2$

$\Rightarrow$ let $SD(X) = \sigma_X$ & $SD(Y) = \sigma_Y$

$$\Rightarrow E(X^*) = E\left(\frac{X-\mu_X}{\sigma_X}\right)$$

$$= \frac{1}{\sigma_x} * E(X - \mu_x)$$

$$= \frac{1}{\sigma_x} * \left(E(X) - E(\mu_x)\right) = \frac{1}{\sigma_x} * (\mu_x - \mu_x)$$

$$= \frac{1}{\sigma_x} * 0$$

$$= 0$$

$\Rightarrow$ for same process $E(Y^*) = 0$

$$\Rightarrow E\left((X^* + Y^*)^2\right) = E\left(X^{*^2} + Y^{*^2} + 2X^*Y^*\right)$$

$$= E(X^{*^2}) + E(Y^{*^2}) + 2E(X^*Y^*).$$

$$\Rightarrow E(X^{*^2}) = E\left(\frac{(X-\mu_x)^2}{\sigma_x^2}\right)$$

$$= \frac{1}{\sigma_x^2} * E(X^2 + \mu_x^2 - 2X\mu_x)$$

$$= \frac{1}{\sigma_x^2} * \left(E(X^2) + E(\mu_x^2) - 2E(X\mu_x)\right)$$

$$= \frac{1}{\sigma_x^2} * \left(E(X^2) + \mu_x^2 - 2\mu_x^2\right) \leftarrow \mu_x \text{ is a constant.}$$

$$= \frac{1}{\sigma_x^2} * \left(E(X^2) - \mu_x^2\right)$$

$$= \frac{1}{\sigma_x^2} * \left(E(X^2) - (E(X))^2\right)$$

$$= \frac{1}{\sigma_x^2} * Var(X) = \frac{1}{\sigma_x^2} * \sigma_x^2$$

$$= 1$$

$\Rightarrow$ for same process $E(Y^{*^2}) = 1$

$$\Rightarrow E\left((X^* + Y^*)^2\right) = E(X^{*^2}) + E(Y^{*^2}) + 2E(X^*Y^*)$$

$$= 2 + 2E(X^* Y^*)$$

$\Rightarrow$ since $(X^* + Y^*)^2 \geq 0$

$\Rightarrow 2 + 2E(X^* Y^*) \geq 0$

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

$$\Rightarrow E((X^* + Y^*)^2) = E(X^{*2}) + E(Y^{*2}) - 2E(X^* Y^*)$$

$$= 2 - 2E(X^* Y^*)$$

$\Rightarrow$ since $(X^* - Y^*)^2 \geq 0$

$\Rightarrow 2 - 2E(X^* Y^*) \geq 0$

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

$\Rightarrow -1 \leq E(X^* Y^*) \leq 1$

$\Rightarrow$ since $E(X^* Y^*) = r(X, Y)$ $\leftarrow$ part a.

$\Rightarrow -1 \leq r(X, Y) \leq 1$

$\Rightarrow$ thus, $-1 \leq r(X, Y) \leq 1$

c. $\rightarrow Y = aX + b$ & $a \neq 0$.

$\rightarrow r(X, Y) = \pm 1$ with $\pm a$

$$\Rightarrow r(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{cov(X, aX + b)}{\sigma_X \sigma_Y}$$

$$= \frac{a * cov(X, X)}{}$$

$$= \frac{\frac{a * Var(X)}{\sigma_X \sigma_Y}}{\sigma_X \sigma_Y}$$

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

$\Rightarrow \sigma_Y = \pm a * \sigma_X$

$$\Rightarrow r(X,Y) = \pm \frac{a * Var(X)}{a * \sigma_X \sigma_X} = \pm \frac{\cancel{a * Var(X)}}{\cancel{a * Var(X)}}$$

$= \pm 1$

$\Rightarrow$ for $a > 0 \rightarrow \sigma_Y = a \sigma_X$

$\Rightarrow$ so $r(X,Y) = 1$

$\Rightarrow$ for $a < 0 \rightarrow \sigma_Y = -a \sigma_X$

$\Rightarrow$ so $r(X,Y) = -1$

$\Rightarrow$ thus, for $Y = aX + b$,

$r(X,Y) = 1$ for positive $a > 0$, and

$r(X,Y) = -1$ for negative $a < 0$.

d. $\rightarrow$ sequence iid. bernoulli (p) trails

$\rightarrow$ postive int K & X_K = # succes in trails $1 \sim K$

$\rightarrow cov(X_n, X_{n+m})$ & bilinearity

$\rightarrow r(X_n, X_{n+m})$

$\Rightarrow$ let I_i be indicator of success for the i^{th} trail

$\Rightarrow X_n = I_1 + I_2 \cdots\cdots I_n$

$\Rightarrow cov(X_n, X_{n+m}) = cov(I_1 + \cdots\cdots I_n, I_1 \cdots\cdots + I_n \cdots\cdots + I_{m+n})$

$\Rightarrow = cov(I_1, I_1 \cdots\cdots I_{n+m}) + \cdots\cdots + cov(I_n, I_1 \cdots\cdots I_{n+m})$.

$= n * cov(I_1, I_1 \cdots\cdots I_{n+m})$

$= n * Var(I_1)$

$= n * pq \leftarrow$ chapter 13

$$\Rightarrow r(X_n, X_{n+m}) = \frac{cov(X_n, X_{n+m})}{\sigma_{x_n} * \sigma_{x_{n+m}}}$$

$$= \frac{n * pq}{\sqrt{npq} * \sqrt{(n+m)pq}}$$

$$= \frac{n}{\sqrt{n} * \sqrt{n+m}}$$

$$= \sqrt{\frac{n}{n+m}}$$

$\Rightarrow$ thus, $r(X_n, X_{n+m}) = \sqrt{\frac{n}{n+m}}$

e.

$\Rightarrow$ as $m \rightarrow \infty$

$\Rightarrow \frac{n}{n+m} \rightarrow$ goes to 0

$\Rightarrow \lim_{m \rightarrow \infty} r(X_n, X_{n+m}) = 0$

$\Rightarrow$ this make sence because as m becomes really large, the result depends more no the n^{th} to $n+m^{th}$ trail and the result of X_n doesn't affect the relation that much, so it make sense that the correlation coeffizient is small.

$\Rightarrow$ thus, the limit of answer to part $d = 0$ as $m \to \infty$ and it is consistent with intuition.