Jack Rodan
STAT 630
Assignment 7
10/29/21

Chap 4 Exercise 4.2.12: [Note. $M_n = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Use R to generate the random variables w/ rexp.] Generate $X_1, \dots, X_n$ iid Exp(5) and compute M_n when $n = 20$. Repeat this $N = 10^5$ times, and compute the proportion of values M_n that lie between 0.19 & 0.21. Repeat this w/ $n = 50$. What property of convergence in probability do your results illustrate?

- $P(0.19 \le M_{20} \le 0.21) \approx 0.1789$, $P(0.19 \le M_{20} \le 0.21) \approx 0.27721$
- This result illustrates the weak LLN.

Exercise 4.4.4: Let W_n have density $\frac{1 + x/n}{1 + 1/2n}$ for $0 < x < 1$ & 0 o.w.
Let $W \sim \text{Unif}[0,1]$. Prove W_n converges in distribution to W.

A [DEF 4.4.1: Let $X_1, X_2, \dots, X_n$ be random variables. Then we say the sequence $\{X_n\}$ converges in distribution to X, if $\forall x \in \mathbb{R}^1$ s.t. $P(X = x) = 0$ we have $\lim_{n \to \infty} P(X_n \le x) = P(X \le x)$
and we write $X_n \xrightarrow{D} X$]

$$\lim_{n \to \infty} W_n = \lim_{n \to \infty} \frac{1 + x/n}{1 + 1/2n} = 1 = W$$

Chap 4 Exercise 4.4.12: Add (d-f): Determine the exact distribution of the average time to service for the first n customers when $n = 16, 36, 100$. Then use pgamma function in R to find the exact probability and compare to the normal approx.

- If the service time, in minutes at a bank has Exp(1/2) distribution. Use the CLT to estimate the probability that the service time of the 1st n customers is less than 2.5 minutes when: (NOTE: let X be the variable of interest)

(a) $n = 16$.

$$P(\bar{X}_{16} \le 2.5) \approx 0.84265$$

(b) $n = 36$

$$P(\bar{X}_{36} \le 2.5) \approx 0.92641$$

(c) $n = 100$

$$P(\bar{X}_{100} \le 2.5) \approx 0.99037$$

3) (Contd)

NOTE: ① IF $Y = \sum_{i=1}^{n} X_i$ where X_i iid $Exp(\lambda)$

then $M_Y(s) = M_{X_1+\dots+X_n}(s) = E[e^{s(X_1+\dots+X_n)}] = \prod_{i=1}^{n} E[e^{sX_i}]$

$= (\lambda/(\lambda-s))^n$

$\Rightarrow \boxed{Y \sim \text{Gamma}(n, \lambda)}$

② $\bar{X} = \frac{1}{n} Y \Rightarrow Y = h^{-1}(\bar{x}) = n\bar{x}$; $\frac{d}{d\bar{x}}[h^{-1}(x)] = n$

$f_{\bar{X}}(\bar{x}) = \left(\frac{(\lambda)^n}{\Gamma(n)}\right)(n\bar{x})^{n-1} e^{-\lambda(n\bar{x})} (n)$

$= \left(\frac{\lambda^n}{\Gamma(n)}\right) n^n \bar{x}^{n-1} e^{-(\lambda n)\bar{x}}$

$f_{\bar{X}}(\bar{x}) = \left(\frac{(\lambda n)^n}{\Gamma(n)}\right) \bar{x}^{n-1} e^{-(\lambda n)\bar{x}}$

$\Rightarrow \bar{X} \sim \text{Gamma}(n, \lambda n)$

(d) $n=16 \Rightarrow \bar{X} \sim \text{Gamma}(16, 8)$

$P(\bar{X} \leq 2.5) = 0.8434869$ which is close to our approximation using the CLT of 0.84265

(e) $n=36 \Rightarrow \bar{X} \sim \text{Gamma}(36, 18)$

$P(\bar{X} \leq 2.5) = 0.9257825$ which is close to our approximation using the CLT of 0.92641.

(d) $n=100 \Rightarrow \bar{X} \sim \text{Gamma}(100, 50)$

$P(\bar{X} \leq 2.5) = 0.9906209$ which is close to our approximation using the CLT of 0.99037

4.) Chp 4 Exercise 4.4.16:

Generate $N = 10^4$ samples of $X_1, X_2, .. X_{30}$ Uniform $[-20, 10]$.
Use these samples to estimate the probability $P(M_{30} \leq -5)$. How does your answer compare to what the CLT gives as an approximation?

- $P(\bar{X} \leq -5) \approx 0.505$ according to our simulation. This is close to the CLT approximation of 0.50.

5.) Chp 4 Exercises 4.6.1, 4.6.2, 4.6.7

Chp 4 Exercise 4.6.1:

Let $X_1 \sim N(3, 2^2)$ and $X_2 \sim N(-8, 5^2)$ be independent. Let $U = X_1 - 5X_2$ and $V = -6X_1 + CX_2$ where C is a constant.

(a) What are the distributions of U & V? (H.0.4 pg 30)*

$U \sim N(43, 629)$

$V \sim N(-18 - 8C, 144 + 25c^2)$

(b) What is the value of C that makes U & V independent? (H.0.4 pg 31)*

NOTE Theorem 4.6.2: If $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, ..., n$ and also that the $\{X_i\}$ are independent. Let $U = \sum_{i=1}^{n} a_i X_i$ and $V = \sum_{i=1}^{n} b_i X_i$ for some constants $\{a_i\}$ & $\{b_i\}$. Then $Cov(U,V) = \sum_{i=1}^{n} a_i b_i \sigma_i^2$. Furthermore, $Cov(U,V) = 0 \iff U$ & V are independent

$$Cov(U,V) = (1)(-6)(2^2) + (-5)(c)(5^2) = 0$$
$$= -24 - 125c = 0$$

(Contd)

Chp4 Exercise 4.6.2: Let $X \sim N(3,5)$, $Y \sim N(-7,2)$ be independent

(a) What's the distribution of $Z = 4X - 1/3 Y$ *(H.O.4 pg 30)

$Z \sim N(4(3) + (-1/3)(-7),\ (4^2)(5) + (-1/3)^2(2))$

$Z \sim N(43/3,\ 722/9)$

(b) What's the cov(X; Z)? (See H.O.3 slide 34) (also Theorem 3.3.2)

Theorem 3.3.2 (Linearity of covariance)

Let X, Y, Z be three random variables. Let $a, b \in \mathbb{R}$. Then

$Cov(aX + bY, Z) = a\,Cov(X,Z) + b\,Cov(Y,Z)$

$Cov(X,Z) = Cov(X, 4X - 1/3 Y) = 4\,cov(X,X) - 1/3\,cov(X,Y)$

* X & Y are ind $\Rightarrow$ cov(X,Y) = 0

$= 4\,var(X)$

$= 4(5) = \boxed{20 = cov(X,Z)}$

Chp 4 Exercise 4.6.7: Let $X_1, \dots, X_n$ be iid $N(0,1)$. Find a value C s.t.

$CX_1 / \sqrt{X_2^2 + \dots + X_{n+1}^2} \sim t(n)$ *(see H.O.4 pg 35)

- From H.O.4 pg 35: If $Z \sim N(0,1)$ & $U \sim \chi^2(n)$ then

$T = Z/(U/n)^{1/2} \sim t(n)$.

$T = \dfrac{\sqrt{n}\,Z}{\sqrt{U}} \Rightarrow \boxed{C = \sqrt{n}}$

Chp 4 Exercise 4.6.10: Let $X_1, \dots, X_{100}$ be iid $N(0,1)$. Compute the dist of the following:

(a) $X_1^2 \sim \chi^2(1)$

(b) $X_3^2 + X_5^2 \sim \chi^2(2)$

(c) $X_{10} / \sqrt{(X_{20}^2 + X_{30}^2 + X_{40}^2)/3} \sim t(3)$

(d) $3X_{10}^2 / [X_{20}^2 + X_{30}^2 + X_{40}^2] \sim F(1,3)$

(e) $30(X_1^2 + \dots + X_{70}^2) / 70(X_{71}^2 + \dots + X_{100}^2) \sim F(70,30)$

7.) β $X_1, \dots X_n$ are i.i.d R.Vs from the Exp(λ) dist. Write down the joint pdf for the random vector $(X_1, \dots, X_n)$.

- $f_{X_i}(x_i) = \lambda e^{-\lambda x_i} \Rightarrow f_{X_1,\dots X_n}(x_1, \dots x_n) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$

$$\boxed{f_{X_1,\dots,X_n}(x_1,\dots x_n) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}}$$

8.) β $T_1, \dots, T_n$ are iid binomial$(4, \theta)$. Write down the pmf for the random vector $(T_1, \dots, T_n)$.

$$f_{T_i}(t_i) = \binom{4}{t_i} \theta^{t_i} (1-\theta)^{4-t_i} = \binom{4}{t_i} \theta^{t_i} (1-\theta)^4 (1-\theta)^{-t_i}$$

$$f_{T_1,\dots T_n}(t_1, \dots, t_n) = \prod_{i=1}^{n} \binom{4}{t_i} \theta^{t_i} (1-\theta)^4 (1-\theta)^{-t_i}$$

$$= \left[\binom{4}{t_1} \cdots \binom{4}{t_n}\right] (1-\theta)^{4n} (1-\theta)^{-\sum_{i=1}^{n} t_i} \theta^{\sum_{i=1}^{n} t_i}$$

$$\boxed{f_{T_1,\dots T_n}(t_1, \dots, t_n) = (1-\theta)^{4n} \left[\prod_{i=1}^{n} \binom{4}{t_i}\right] \left[\left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n} t_i}\right]}$$

9.) Let $X_1, \dots, X_n$ be iid w/ $f_{X_i}(x_i) = \frac{1}{2\beta} e^{-|x_i - \mu|/\beta}$. Write down the joint pdf of the random vector $(X_1, \dots, X_n)$. Simplify as possible.

$$f_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{2\beta} e^{-|x_i - \mu|/\beta}$$

$$= \left(\frac{1}{2\beta}\right)^n e^{-\left(\frac{1}{\beta}\right) \sum_{i=1}^{n} |x_i - \mu|}$$

$$\boxed{f_{X_1,\dots,X_n}(x_1, \dots, x_n) = \left(\frac{1}{2\beta}\right)^n \sum_{k=0}^{\infty} \frac{-(1/\beta)^k \left(\sum_{i=1}^{n} |x_i - \mu|\right)^k}{k!}}$$

10.) Use R to simulate $N = 10^4$ random samples $(Z_1, \dots, Z_n)$ from a $N(0,1)$, and compute $T = \max(Z_1, \dots, Z_n)$ for each sample. Use $n = 20$. Plot a histogram and comment on its shape and symmetry.

- The histogram of T looks close to a normal distribution for $n=20$. However it is skewed slightly to the right. As n gets larger the right skew becomes more prominent.