Exercise2

ZHIYU LI, 202104615

October 2021

1

See in the last 4 pages

2

(a^*) Cauchy

$F_X(x) = \int f_X(x)dx = \int \frac{1}{\pi(1+x^2)}$dx

Let x = tanθ , then:

F(θ) = $\int \frac{1}{\pi(1+\frac{sin\theta^2}{cos\theta^2})} dtan\theta$

= $\frac{1}{\pi}\int \frac{1}{1+\frac{sin\theta^2}{cos\theta^2}} \frac{1}{cos\theta^2}d\theta$

=$\frac{1}{\pi}\int 1d\theta$

= $\frac{1}{\pi}$(θ+C)

As θ = arctan(x), then:

$F_X(x)$ = $\frac{1}{\pi}$(arctan(x)+C)

According to the property of CDF, $\lim_{x\to\infty}$F(x) = 1.

As $\lim_{x\to\infty}$arctan(x) = $\frac{\pi}{2}$, then C = $\frac{\pi}{2}$.

$F_X(x) = \frac{arctan(x)}{\pi}+\frac{1}{2}$

(b^*) Logistic

$F_X(x) = \int f_X(x)dx = \int \frac{e^{-x}}{(1+e^{-x})^2}$dx

$=\frac{1}{1+e^{-x}}$+C

As $\lim_{x\to\infty}\frac{1}{1+e^{-x}} = 1$, then C = 0.

$F_X(x) = \frac{1}{1+e^{-x}}$

(c^*) Pareto

$F_X(x) = \int f_X(x)dx = \int \frac{a-1}{(1+x)^a}$dx

$= -(1+x)^{1-a}$ + C

As $\lim_{x\to\infty} -(1+x)^{1-a} = 0$, then C = 1.

$F_X(x) = -(1+x)^{1-a} + 1 \quad (x > 0)$

$F_X(x) = 0 \quad (x <= 0)$

(d^*) Weibull

$F_X(x) = \int f_X(x)dx = \int c^{\tau-1}e^{-cx^\tau}dx$

$= -e^{-cx^\tau}$ + C

As $\lim_{x\to\infty} -e^{-cx^\tau} = 0$, then C = 1.

$F_X(x) = -e^{-cx^\tau} + 1 \quad (x > 0)$

$F_X(x) = 0 \quad (x <= 0)$

3

(a^*) Gamma Distribution

Xs from Gamma Distribution are continuous. Therefore, $E(X) = \int xf_X(x)dx$.

$\int_{-\infty}^{\infty} xf_X(x)dx = \int_0^\infty x\frac{e^{-kx}x^{r-1}k^r}{(r-1)!}dx + \int_{-\infty}^0 0dx = \int_0^\infty \frac{e^{-kx}x^{r-1}k^r}{(r-1)!}dx$

$= \int_0^\infty x\frac{e^{-kx}x^{r-1}k^r}{(r-1)!}dx = \int_0^\infty \frac{e^{-kx}x^r k^r}{(r-1)!}dx = \frac{r}{k}\int_0^\infty \frac{e^{-kx}x^r k^{r+1}}{(r)!}dx = \frac{r}{k}$

E(X)=$\frac{r}{k}$

Var(X)=$E(x^2)-E(x)^2$

$E(X^2)=\int_{-\infty}^{\infty}x^2f_X(x)dx$

Likewise, $E(X^2)=\frac{r^2+r}{k^2}$. Var(X)=$\frac{r}{k^2}$

E(X)=$\frac{r}{k}$, Var(X)=$\frac{r}{k^2}$

($b*$) Poisson Distribution

E(X)=$\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{(x-1)!}$=$\lambda\sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}+0$=$\lambda$

$E(X^2)=\sum_{x=0}^{\infty}\frac{(x-1)e^{-\lambda}\lambda^x+e^{-\lambda}\lambda^x}{(x-1)!}$=$\lambda^2$+$\lambda$

Var(X)=λ

(c*)Pareto Distribution

E(X)=$\int_0^{\infty}\frac{x(a-1)}{(1+x)^a}dx$=$\int_0^{\infty}\frac{x(a-1)+(a-1)-(a-1)}{(1+x)^a}dx$=

=$\int_0^{\infty}\frac{(a-1)}{(1+x)^{a-1}}-\frac{a-1}{(1+x)^a}dx$

=$\frac{a-1}{2-a}(1+x)^{2-a}|_0^{\infty}$+$(1+x)^{1-a}|_0^{\infty}$=$\frac{1}{a-2}$

$E(X^2)$=$\int_0^{\infty}\frac{a-1}{(1+x)^{a-2}}-\frac{2(a-1)}{(1+x)^{a-1}}+\frac{a-1}{(1+x)^a}dx$

=$[\frac{a-1}{3-a}(1+x)^{3-a}-\frac{2(a-1)}{2-a}(1+x)^{2-a}-(1+x)^{1-a}]|_0^{\infty}$=$\frac{2}{(3-a)(2-a)}$

Var(X)=$\frac{1-a}{(3-a)(2-a)^2}$

(d)Negative Binomial

$P(X=x)=\frac{(a+x-1)!}{x!(a-1)!}[\frac{b}{1+b}]^a[\frac{1}{1+B}]^x$

For convenience, I substitute $\frac{b}{1+b}$ with p, then the equation above can be expressed as:

$P(X=x)=C_{a+x-1}^{x}p^a(1-p)^x$

$E(X)=\sum_0^{\infty}xP(x)=\sum_1^{\infty}\frac{(a+x-1)!}{(x-1)!(a-1)!}p^a(1-p)^x$

(As $0*P(0)=0$, summation starting from is equivalent to the original one)

$\sum_{x=1}^{\infty} \frac{(a+x-1)!}{(x-1)!(a-1)!} p^a(1-p)^x = \sum_{x=1}^{\infty} aC_{a+x-1}^{x-1} p^a(1-p)^x$

Let t=x-1, then

$\sum_{1}^{\infty} aC_{a+x-1}^{x-1} p^a(1-p)^x = \sum_{t=0}^{\infty} aC_{a+t}^{t} p^a(1-p)^{t+1}$

$= \frac{a(1-p)}{p} \sum_{t=0}^{\infty} C_{a+t}^{t} p^{a+1}(1-p)^t$

$C_{a+t}^{t} p^{a+1}(1-p)^t$ is a density function of negative binomial. Hence, $\sum_{t=0}^{\infty} C_{a+t}^{t} p^{a+1}(1-p)^t = 1$.

Finally, $E(X) = \frac{a(1-p)}{p} = \frac{a}{b}$

$Var(X) = E(X^2) - E(X)^2$

$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{(a+x-1)!}{x!(a-1)!} p^a(1-p)^x$

$= \sum_{x=1}^{\infty} \frac{x(a+x-1)!}{(x-1)!(a-1)!} p^a(1-p)^x$

$= \sum_{x=1}^{\infty} axC_{a+x-1}^{x-1} p^a(1-p)^x$

Let t=x-1, then

$\sum_{x=1}^{\infty} axC_{a+x-1}^{x-1} p^a(1-p)^x = \sum_{t=0}^{\infty} a(t+1)C_{a+t}^{t} p^a(1-p)^{t+1}$

$= \sum_{t=0}^{\infty} \frac{a(t+1)(1-p)}{p} C_{a+t}^{t} p^{a+1}(1-p)^t$

$= \frac{a(1-p)}{p} \sum_{t=0}^{\infty} (t+1)C_{a+t}^{t} p^{a+1}(1-p)^t$

$\sum_{t=0}^{\infty} (t+1)C_{a+t}^{t} p^{a+1}(1-p)^t = \sum_{t=0}^{\infty} C_{a+t}^{t} p^{a+1}(1-p)^t + \sum_{t=0}^{\infty} tC_{a+t}^{t} p^{a+1}(1-p)^t$

$\sum_{t=0}^{\infty} C_{a+t}^{t} p^{a+1}(1-p)^t = 1$

$\sum_{t=0}^{\infty} tC_{a+t}^{t} p^{a+1}(1-p)^t = \frac{(1+p)(a+1)}{p}$

(The calculating process is slightly omitted here, as the procedure is almost the same as that in calculating E(X))

Just combine the two parts of $\sum_{t=0}^{\infty} (t+1)C_{a+t}^{t} p^{a+1}(1-p)^t$,

$E(X^2) = \frac{a(1-p)}{p} \sum_{t=0}^{\infty} (t+1)C_{a+t}^{t} p^{a+1}(1-p)^t = \frac{a(1-p)}{p}\left(\frac{(1-p)(a+1)}{p} + 1\right)$

$Var(X) = E(x^2) - E(X)^2 = \frac{a(1-p)}{p}\left(\frac{(1-p)(a+1)}{p} + 1\right) - \left[\frac{a(1-p)}{p}\right]^2 = \frac{a(1-p)}{p^2}$

Replace p with the original expression $\frac{b}{1+b}$,

$Var(X) = \frac{a(b+1)}{b^2}$

$E(X) = \frac{a}{b}$, $Var(X) = \frac{a(b+1)}{b^2}$.

4

Exponential Distribution

As both of x and y are greater than 0, then $P(x + y > X, X > x) = P(x + y > X)$.

$P(x + y > X|x > X) = \frac{P(x+y>X)}{P(x>X)}$.

$P(x > X) = 1 - \int_0^x \frac{1}{\lambda} e^{\frac{-x}{\lambda}} dx = \int_x^\infty \frac{1}{\lambda} e^{\frac{-x}{\lambda}} dx$

$= -e^{\frac{-x}{\lambda}} |_x^\infty = e^{\frac{-x}{\lambda}}$

Likewise, $P(y > X) = e^{\frac{-y}{\lambda}}$, and $P(x + y > X) = e^{\frac{-(x+y)}{\lambda}}$

$P(x + y > X|x > X) = \frac{P(x+y>X)}{P(x>X)} = P(y > X) = e^{\frac{-y}{\lambda}}$

Geometric Distribution

$P(x > X) = 1 - \sum_1^x q^{x-1} p$

$\sum_1^x q^{x-1} p$ is a geometric sequence whose sum equals to $\frac{(q^x-1)p}{q-1}$.

$P(x > X) = 1 - \frac{(q^x-1)p}{q-1}$

Likewise, $P(y > X) = 1 - \frac{(q^y-1)p}{q-1}$, and $P(x + y > X) = 1 - \frac{(q^{x+y}-1)p}{q-1}$

Substitute p with 1-q, then we can get:

$P(x > X) = q^x, P(y > X) = q^y, and P(x + y > X) = q^{x+y}$.

It is obvious that $P(x + y > X|x > X) = \frac{P(x+y>X)}{P(x>X)} = P(y > X) = q^y$

So far, I have shown that if X has either the exponential distribution, or a geometric distribution, then X has no memory.

"No memory" means what happens in the process will not affect the probability of the next event. To draw an economics analogy, suck cost should not affect the future decision, as it does not matter after happening.

5

(a)

As Y = e^X, then $X = lnY$.

$F_X(x) = F_X(lny)$

$f(y)=\frac{\partial F(lny)}{\partial y}=\frac{f_X(lny)}{y}=\frac{1}{y\sqrt{2\pi\sigma}}\exp(-\frac{(lny-\mu)^2}{2\sigma^2})$

$E(Y^n) = E(e^{nx}) = \int e^{xn}f(x)dx$

$=\int e^{nx}\frac{1}{\sqrt{2\pi\sigma}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}dx$

$=\int \frac{1}{\sqrt{2\pi\sigma}}e^{\frac{2\sigma^2 nx-x^2-\mu^2+2\mu x}{2\sigma^2}}dx$

$=\int \frac{1}{\sqrt{2\pi\sigma}}e^{\frac{(\sigma^2 n+\mu)^2-(x-\sigma^2 n-\mu)^2-\mu^2}{2\sigma^2}}dx$

$=\frac{(\sigma^2 n+\mu)^2-\mu^2}{e^{2\sigma^2}}\int \frac{1}{\sqrt{2\pi\sigma}}e^{\frac{-(x-\sigma^2 n-\mu)^2}{2\sigma^2}}dx$

$\int \frac{1}{\sqrt{2\pi\sigma}}e^{\frac{-(x-\sigma^2 n-\mu)^2}{2\sigma^2}}dx=1$

Therefore, $E(Y^n) = e^{\frac{\sigma^4 n^2+2\sigma^2 n\mu}{2\sigma^2}}$

$E(Y) = e^{\frac{\sigma^2}{2}+\mu}$, $E(Y^2) = e^{2\sigma^2+2\mu}$

$Var(Y) = E(X^2) - E(Y)^2 = e^{\sigma^2+2\mu}(e^{\sigma^2} - 1)$.

Finally, $E(Y) = e^{\frac{\sigma^2}{2}+\mu}$, $Var(Y) = e^{\sigma^2+2\mu}(e^{\sigma^2} - 1)$.

(b) See in the last 4 pages

RExersice_W2

Zhiyu Li

2021/10/10

Q1

```r
#Load necessary package(s)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(ggplot2)
library(ggpubr)
invisible()
```

```r
#Define a function to generate n times tossing
tossing <- function(n){
  return(ceiling(runif(n, 0, 6)))
}
```

```r
#Define a function for calculating P(A),P(B),P(AB)
Prob_A_B_AB <- function(n, A, B){
  #assign tossing result to a local variable
  n_times_tossing <- tossing(n)

  Prob_A <- sum(n_times_tossing %in% A)/n

  Prob_B <- sum(n_times_tossing %in% B)/n

  Prob_AB <- sum((n_times_tossing %in% intersect(A, B)))/n

  return(c(Prob_A, Prob_B, Prob_AB))
}
```

```r
#Define a function for calculating the difference between
#P(A)*P(B) and P(AB)
Diff_Between_A_B_AB <- function(n, A = c(2,4,6), B = c(1,2,3,4)){
  Prob <- Prob_A_B_AB(n, A, B)
  ProbA <- Prob[1]
  ProbB <- Prob[2]
  ProbAB <- Prob[3]
  Diff <- ProbA * ProbB - ProbAB
  return(Diff)
}
```

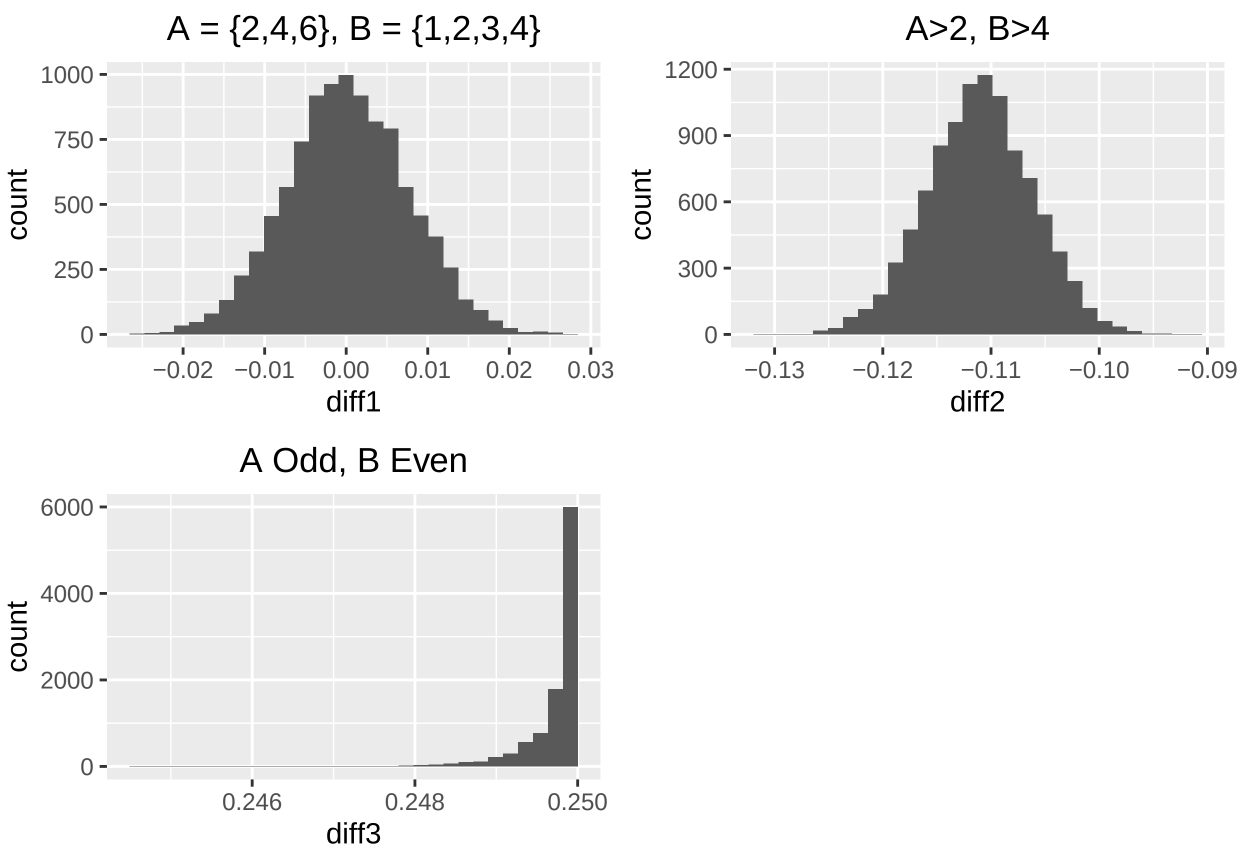
```r
#Default setting is A = {2,4,6}, B = {1,2,3,4}
diff1 <- replicate(10000, Diff_Between_A_B_AB(1000))
diff1_df <- as.data.frame(diff1)
d1<- ggplot(data = diff1_df, aes(x = diff1))+
  geom_histogram()+
  labs(title = 'A = {2,4,6}, B = {1,2,3,4}')+
  theme(plot.title = element_text(hjust = 0.5))

#Let A > 2, that A = {3,4,5,6}, and B > 4, that B = {5,6}.
#A and B are not independent.
A = c(3,4,5,6)
B = c(5,6)
diff2 <- replicate(10000, Diff_Between_A_B_AB(1000, A, B))
diff2_df <- as.data.frame(diff2)
d2<- ggplot(data = diff2_df, aes(x = diff2))+
  geom_histogram()+
  labs(title = 'A>2, B>4')+
  theme(plot.title = element_text(hjust = 0.5))

#Let A is odd, that A = {1,3,5}, and B is even, that B = {2,4,6}.
#A and B are not independent.
A = c(1,3,5)
B = c(2,4,6)
diff3 <- replicate(10000, Diff_Between_A_B_AB(1000, A, B))
diff3_df <- as.data.frame(diff3)
d3<- ggplot(data = diff3_df, aes(x = diff3))+
  geom_histogram()+
  labs(title = 'A Odd, B Even')+
  theme(plot.title = element_text(hjust = 0.5))

ggarrange(d1,d2,d3)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
#Means for Differences of 3 Events
c(
  Independent_Events = diff1 %>% mean(),
  Dependent_Events_I = diff2 %>% mean(),
  Dependent_Events_II = diff3 %>% mean()
)
```

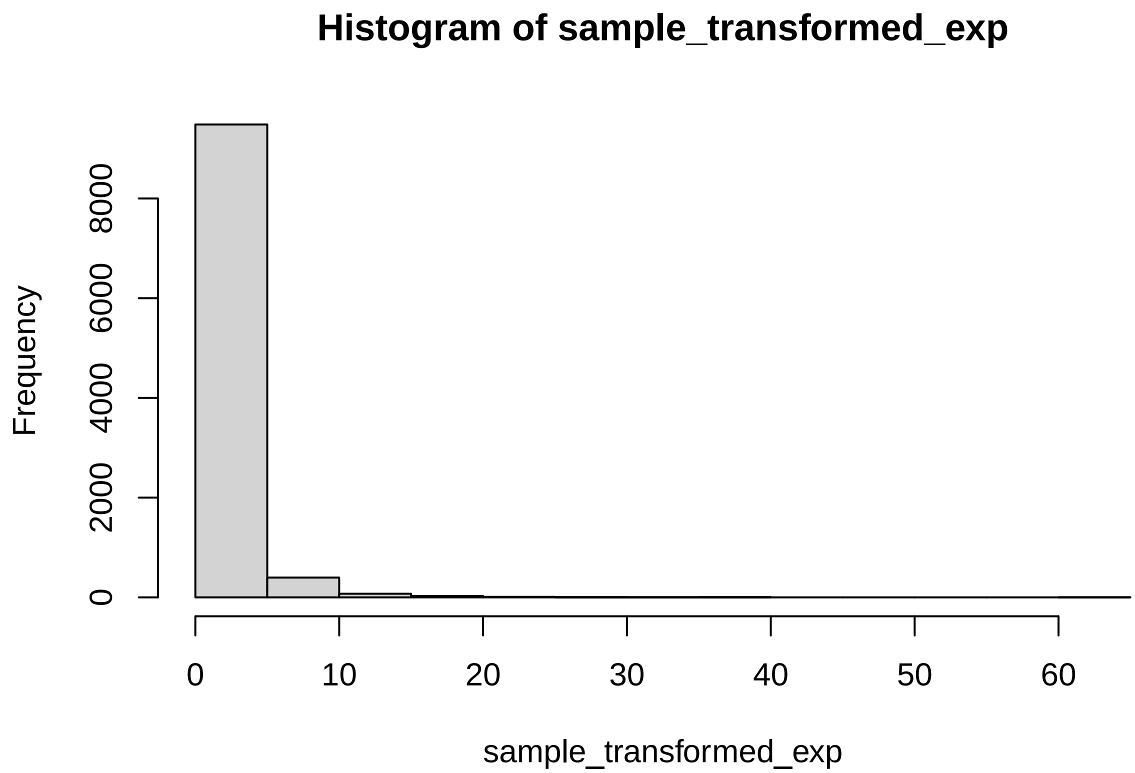
```
##  Independent_Events  Dependent_Events_I Dependent_Events_II
##        0.0001266412       -0.1110667955        0.2497407580
```

Q5(b)

```
#Draw a sample of 10,000 from N(0,1)
sample_normal <- rnorm(10000, mean = 0, sd = 1)
sample_transformed_exp <- exp(sample_normal)
mean_sample_transformed <- mean(sample_transformed_exp)
sd2_sample_transformed <- sd(sample_transformed_exp)^2
cat(' sample mean:', mean_sample_transformed,'\n','sample variance',sd2_sample_transformed)
```

```
##  sample mean: 1.651046
##  sample variance 5.21628
```

```r
#Histogram for Y
hist(sample_transformed_exp)
```




```r
#Define function for calculating mean of y
mean_of_y <- function(mu, sigma){
  mean <- exp((sigma^2)/2 + mu)
  return(mean)
}
#Define function for calculating variance of y
sd_square_of_y <- function(mu, sigma){
  std2 <- exp(sigma^2 + 2*mu)*(exp(sigma^2)-1)
  return(std2)
}
```

```r
cat(' sample mean,expectation',mean_sample_transformed,mean_of_y(0,1),'\n','sample variance,Variance',sd
```

```
##  sample mean,expectation 1.651046 1.648721
##  sample variance,Variance 5.21628 4.670774
```