STAT3701HW1

Mingming Xu

6/21/2019

Problem 1:

Some Basic Notes on Probability:

Excersice 1:

ii.If A and B are independent,then A,$B'$are independent.

Prove:

Because A and B are independent, $P(A \cap B) = P(A) * P(B)$

$A = (A \cap B) \cup (A \cap B')$

$P(A) = P((A \cap B) \cup (A \cap B')) = P(A \cap B) + P(A \cap B') = P(A) * P(B) + P(A \cap B')$

So, $P(A \cap B') = P(A) - P(A) * P(B) = P(A) * (1 - P(B)) = P(A) * P(B')$, which implies A,$B'$are independent.

iii. If A and B are independent,then A',B are independent.

Prove:

Because A and B are independent, $P(A \cap B) = P(A) * P(B)$

$B = (B \cap A) \cup (B \cap A')$

$P(B) = P((B \cap A) \cup (B \cap A')) = P(B \cap A) + P(B \cap A') = P(B) * P(A) + P(B \cap A')$

So, $P(B \cap A') = P(B) - P(B) * P(A) = P(B) * (1 - P(A)) = P(B) * P(A')$, which implies A',B are independent.

iv: If A and B are independent,then A',B' are independent.

Prove:

Because A and B are independent, $P(A \cap B) = P(A) * P(B)$

Because $(A' \cap B')$=$(A \cup B)'$, $P(A' \cap B') = P(A \cup B)' = 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B)) = 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A) - P(B) + P(A) * P(B) = (1 - P(A)) * (1 - P(B)) = P(A') * P(B')$, which implies A',B' are independent.

Exercise 3: Show that $P(A) = P(A \mid B)P(B) + P(A \mid B')P(B')$ when 0 < P(B) <1.

Prove: When 0 < P(B) <1,

$P(A \mid B)P(B) + P(A \mid B')P(B') = (P(A \cap B)/P(B))P(B) + (P(A \cap B')/P(B'))P(B') = P(A \cap B) + P(A \cap B') = P(A)$ because $A = (A \cap B) \cup (A \cap B')$

Exercise 4: $var(X) = E(X^2) - E(X)^2$

Prove: $var(X) = E[(X - E(X))^2] = E[X^2 - 2E(X)X + E(X)^2] = E(X^2) - 2E(X)E(X) + E(X)^2 = E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2$

Problem 2:

Some known distributions and other stuff:

Exercise 3: Calculate the mean X, where X ~ geomeric(θ), $\theta \in [0, 1]$

P(X=x)=$\theta(1 - \theta)^{x-1}$

E(X)=$\sum x\theta(1-\theta)^{x-1}$, x=1,2,3. . . n

E(X)=$\theta\sum x(1-\theta)^{x-1}$,x=1,2,3. . . n

let $S=\sum x(1-\theta)^{x-1}$,x=1,2,3. . . n,and let q be $1-\theta$

$S=1+2q+3q^2+...+nq^{n-1}$

$qS=1q+2q^2+3q^3+...+nq^n$

$S-qS=1+q+q^2+q^3+...+q^{n-1}-nq^n$

$S(1-q)=((1-q^n)/(1-q))-nq^n$

$S=(1-q^n)/(1-q)^2-nq^n/(1-q)$, $\theta\in[0,1]$ and $q=1-\theta$ so that $q\in[0,1]$

$\lim_{n\to\infty}S=1/(1-q)^2$ because of $\lim_{n\to\infty}q^n$=0

$\lim_{n\to\infty}S=1/\theta^2$

E(X)=$\theta*1/\theta^2$=$1/\theta$

Excersice 5:

The distribution function $F(x,\theta_1,\theta_2)$:

$F(x,\theta_1,\theta_2)$=0, if $x<\theta_1$

$F(x,\theta_1,\theta_2)$=1, if $x>\theta_2$

$F(x,\theta_1,\theta_2)$=$\int_{\theta_1}^{x}1/(\theta_2-\theta_1)\,du$=$(x-\theta_1)*(1/(\theta_2-\theta_1))$, if $\theta_1\le x\le\theta_2$

Exercise 6: Suppose you are given U ~$U(0,1)$. From this how you generate $U\sim U(a,b)$. Implememt this in R.

```
unif.gen=function(n,a,b){
  u=runif(n)
  u1=(b-a)*u+a

  return (u1)
}
```

Problem 3:

U_1~$U(a,b)$, the pdf $f(u_1,a,b)=1/(b-ab)I_{(a<u_1<b)}$ and the expectation is $E(U_1)=\int_a^b(1/(b-a))*u_1\,dx==(a+b)/2$

when a=1 and b=3 ,the mean of the uniform distribution with those parameters is $(1+3)/2=2$

```
mean(unif.gen(1000,1,3))
```

```
## [1] 2.011866
```

The result shows that the sample mean and the population mean are close.Because the sample is taken from the whole population, so it can reflect the population's characteristics, and the average value of the sample will be closer to the population mean.

Problem 4 :

Use the function given in class to simulate 100 observation fromBer(.5) andBer(.9). Calculate the standard deviationof both samples. Which has more standard deviation and why?(set seed 1729 for this problem.)

```
set.seed(1729)
x1=rbinom(100,1,.5)
x2=rbinom(100,1,.9)
sd(x1)
```

```
## [1] 0.5025189
```

```r
sd(x2)
```

```
## [1] 0.2386833
```

The standard deviation of samples from Ber(0.5) is larger. Because the $sd^2 = p(1-p)$ which is a quadratic image, sd has the largest number when p=0.5.And, sd is increasing when 0<p<0.5 and decreasing when 0.5<p<1.

Problem 5:

i. p(on each pull you reel in a fish that you can take home)=p(W>32)=1-p(w<=32)

```r
1-pnorm(32,29,sqrt(5))
```

```
## [1] 0.08985625
```

ii.

```r
gamma=1-pnorm(32,29,sqrt(5))
E_G=1/gamma
E_G
```

```
## [1] 11.12889
```

The expectation of G is $1/\gamma$,11.12889.

iii.

```r
set.seed(3701)
weigh=0
p=0

while (weigh<=32) {
  weigh=rnorm(1,29,sqrt(5))
 p=p+1
}

p
```

```
## [1] 6
```

Problem 6:

The joint probability distribution $F_{X,Y}(x,y) = P(X \le x, Y \le y)$

i. If X, Y are independent, then $F_{X,Y}(x,y) = P(X \le x, Y \le y) = P(X \le x)P(Y \le y) = F_X(x)F_Y(y)$

ii. If X, Y are independent, $f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_X(x)F_Y(y) = \frac{\partial}{\partial x} F_X(x) \frac{\partial}{\partial y} F_Y(y) = f_X(x)f_Y(y)$

iii.If X, Y are independent:

For discrete random variables X and Y,

$E(XY) = \sum_i \sum_j x_i y_j f_{X_i,Y_j}(x,y) = \sum_i \sum_j x_i y_j f_{X_i}(x) f_{Y_j}(y) = \sum_i x_i f_{X_i}(x_i) \sum_j y_j f_{Y_j}(y_j) = E(X)E(Y)$

For continuous random variables X and Y:

$E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{X,Y}(x,y)\, dx\, dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_X(x) f_Y(y)\, dx\, dy = \int_{-\infty}^{+\infty} y f_Y(y)\, dy \int_{-\infty}^{+\infty} x f_X(x)\, dx = E(X)E(Y)$

iv.If X,Y are independent,$Cov(X,Y) = E\{(X - E(X))(Y - E(Y))\} = E\{XY - XE(Y) - YE(X) + E(X)E(Y)\} = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) = E(X)E(Y) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) = 0$

v.

```
set.seed(1729)
y1=rbinom(1000,1,.5)
y2=rnorm(1000,0,1)
cov(y1,y2)
```

```
## [1] -0.01611453
```

The covariance is close to 0.

Problem 7:

Consider the region given by x=0,x=1,y=0,y=1, which is a square with tops (0,0),(1,0),(0,1),(1,1). And, sketch a circle with center at (0.5,0.5) and radius 0.5 is inside this square.

Then for U_1, U_2 iid $U(0,1)$, P(the point (U_1, U_2) falls in this circle)=the area of the circle/the area of the square=0.25π. The point (U_1, U_2) falls in this circle if $((u_1 - 0.5)^2 + (u_2 - 0.5)^2) \leq 0.5^2$.

```
##generate x_1,x_2...x_n iid ~U(0,1)
x.list=runif(1000000)
##generate y_1,y_2...y_n iid ~U(0,1)
y.list=runif(1000000)
##calculate the proportion of these points falling in the circle
##p=mean(1*(x.list-0.5)^2+(y.list-0.5)^2<=0.25)
pro=ifelse((x.list-0.5)^2+(y.list-0.5)^2<=0.25,1,0)
p=mean(pro)

##pi is equal to 4 times p
pi=4*p
pi
```

```
## [1] 3.1408
```

Problem 8:

If X ~Cauchy(0,1), the pdf $f_X(x;0,1) = \frac{1}{\pi(1+x^2)}$.

The distribution function $F_X(x;0,1) = \int_{-\infty}^{x} \frac{1}{\pi(1+u^2)}\,du = \frac{1}{\pi} arctan(u)|_{-\infty}^{x} = \frac{1}{\pi} arctan(x) + \frac{1}{2}$

Suppose u~unifor(0,1) and let $u = F_X(x)$, then $\pi(u - \frac{1}{2}) = arctan(x)$, then $x = tan((u - \frac{1}{2})\pi)$.

```
stanCau.gen=function(n){
u.list=runif(n)
x.list=tan((u.list-1/2)*pi)
return (x.list)
}
set.seed(1729)
```

Problem 9:

Let $f_X(x) = \int_{-1}^{1} \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}}\,dx$, we set $x = siny$, then $dx = cosydy$.

So, $1 - x^2 = 1 - sin^2 y = cos^2 y$. Also,because x is from -1 to 1 such that y is from $\frac{-\pi}{2}$ to $\frac{\pi}{2}$.

$\int_{-1}^{1} \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}}\,dx$

$=\int_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\pi} \frac{1}{cosy} cosy\,dy$

$=\frac{1}{\pi} y|_{\frac{-\pi}{2}}^{\frac{\pi}{2}}$

$=\frac{1}{\pi}(\frac{\pi}{2}-(-\frac{\pi}{2}))$

$=1$

Because $\int_{-1}^{1}\frac{1}{\sqrt{1-x^2}}\,dx=\pi$.By Monte Carlo Integration, if we take $u_1,u_2...u_n$ iid $\sim uniform(-1,1)$,$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{(1-u_i^2)}}$ $*\,(1-(-1))$is the approxiation of this integration.

```
n=100000
u=unif.gen(n,-1,1)
est=mean((1/(sqrt(1-u^2))))
(1-(-1))*est
```

```
## [1] 3.15331
```