Homework 2

Chenxupeng 2014012882

2017 年 10 月 3 日

1 Exercise 4.2 in Chapter 2

according to the definition of independence:

A is independent of itself $\Leftrightarrow P(A)\bullet P(A) = P(A \cap A) \Leftrightarrow (P(A))^2 = P(A) \Leftrightarrow P(A) = 0$ or 1

2 Exercise 2.9 in Chapter 4

(i) $f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \int_0^x ye^{-x}dy = \frac{e^{-x}x^2}{2}$

$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = \int_y^{+\infty} ye^{-x}dx = ye^{-y}$

we can have that $f_X(x) = \begin{cases} \frac{e^{-x}x^2}{2} & 0 < y \le x \\ 0 & \text{otherwise} \end{cases}$ and $f_Y(y) = \begin{cases} ye^{-y} & 0 < y \le x \\ 0 & \text{otherwise} \end{cases}$

(ii)According to the definition of conditional probablity

$$f_{X|Y}(\bullet | y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} e^{-x+y} & 0 < y \le x \\ 0 & otherwise \end{cases}$$

$$f_{Y|X}(\bullet | x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \frac{2y^2}{x^2} & 0 < y \le x \\ 0 & otherwise \end{cases}$$

(iii)because y = ln2 < 2ln2, so:

$$P(X > 2\ln 2 \,|\, y = \ln 2) = \int_{2\ln 2}^{+\infty} e^{-x+\ln 2}dx = \frac{1}{2}$$

3 Exercise 1.7 in Chapter 5

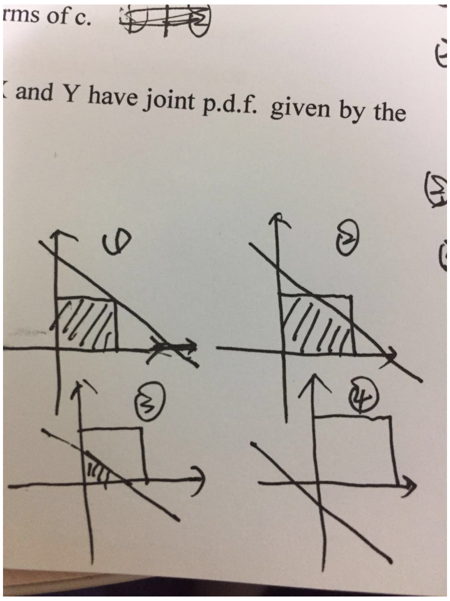
(i) $f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \int_0^1 dy = 1$

$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = \int_0^1 dx = 1$

(ii)From (i) we now that

$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$ for $0 < x, y < 1$

(iii) we can calculate the probability by drawing each conditions:



$$P(X+Y<c) = \begin{cases} 0 & c<0 \\ \dfrac{c^2}{2} & 0 \le c < 1 \\ 1-\dfrac{(2-c)^2}{2} & 1 \le c < 2 \\ 1 & c \ge 2 \end{cases}$$

(iv)

when $c = \frac{1}{4}$, it matches the second condition,so $P(X+Y<c)=\frac{1}{2\cdot 4^2}=\frac{1}{32}$

4 Exercise 1.11 in Chapter 5

(i)

$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \int_{-\infty}^{+\infty} g(x)h(y)dy = g(x)\int_{-\infty}^{+\infty} h(y)dy$ =cg(x)

$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \int_{-\infty}^{+\infty} g(x)h(y)dx = h(y)\int_{-\infty}^{+\infty} g(x)dx$ $=\frac{1}{c}$h(y)

(ii)

$f_X(x) \cdot f_Y(y) = g(x)h(y)(\int_{-\infty}^{+\infty} h(y)dy)\cdot(\int_{-\infty}^{+\infty} g(x)dx)$ = cg(x)* $\frac{1}{c}$h(y) = g(x)h(y)

So $f_{X,Y}(x,y)$=$g(x)h(y)$=$f_X(x)\cdot f_Y(y)$, so that X and Y are independent.

(iii)

If $h = g$, then$\int_{-\infty}^{\infty} g(x)dx$=$\frac{1}{c}$=$\int_{-\infty}^{\infty} h(x)dx$=$\int_{-\infty}^{\infty} h(y)dy = c$,so c = 1/-

1, so we have that: $f_X(x)=f_Y(y)$, so X and Y are identically distributed.

(iv) X and Y are identically distributed, we can say that $P(X>Y)=P(X<Y)$

we have that $P(X>Y)+P(X<Y)+P(X=Y)=1$

and of course P(X=Y) = 0

So $P(X>Y)=\frac{1}{2}$

5 Exercise 1.13 in Chapter 5

(i) The p.d.f. of X is $f(x)=1(0<x<1)$

For a which is positive:

$P(y<a)=P(-\ln x<a)=P(x>e^{-a})$=1-$F(e^{-a})$=1-e^{-a}

So we can have the d.f. of Y is $F_Y(y)=1-e^{-y}\quad(0<y<+\infty)$

And we can have Y's p.d.f. is $f_Y(y)=e^{-y}(0<y<+\infty)$

(ii)

For exponential distribution which λ = 1 The MGF of each Y_i is $M_{Y_i}(t)=\frac{1}{1-t}$

So the MGF of the sum is $M_Z(t)=\frac{1}{(1-t)^n}$, it is the form of Gamma Distribution and the parametres α=$n,\beta=1$. So the p.d.f. should be $F_Z(x)=\frac{x^{n-1}e^{-x}}{(n-1)!}$,x>0.

6 Exercise 1.10 in Chapter 7

According to CLT theorem, $\frac{\sqrt{n}\sum_{i=1}^{n}X_i}{n}\xrightarrow{\text{d}}Z\sim N(0,1)$ (1)

According to WLLN theorem, $\frac{\sum_{i=1}^{n} X_i^2}{n} \xrightarrow{P} EX_i^2 = \sigma^2(X_i) = 1$ (2)

According to Slutsky Theorem, when n approaches ∞ the ratio of equation (1) and (2) approaches the normal distribution $Z \sim N(0,1)$

Similarly, according to WLLN, $\sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n}} \xrightarrow{P} s.d.(X_i) = 1$ (3)

So the ratio of (1) and (3) approaches the normal distribution $Z \sim N(0,1)$

7.

(i) The sample space is $X=(x_1, x_2, x_3, x_4, x_5)$, for $x_i = 0$ or 1

The distribution function is $F(x_i) = \begin{cases} 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \\ 0 & otherwise \end{cases}$ i=1,2,3,4,5

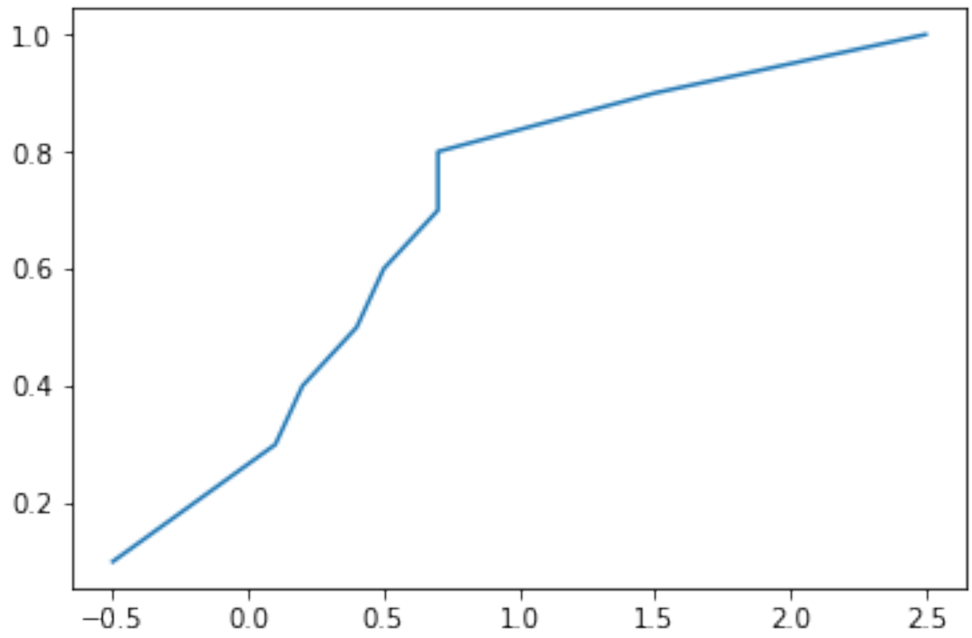
(ii)the last three have the definite numerical values, so the first two are statistics.

(iii)The empirical distribution function is $F(x_i) = \begin{cases} \frac{n-m}{n} & 0 \le x < 1 \\ 1 & x \ge 1 \\ 0 & otherwise \end{cases}$

8.python codes:

```
from statsmodels.distributions import ECDF
import matplotlib.pyplot as plt
a = [-0.5,-0.2,0.1,0.2,0.4,0.5,0.7,0.7,1.5,2.5]
ecdf = ECDF(a)
plt.plot(ecdf.x, ecdf.y)
```

plt.show()



9.

The sample size is n, then $\overline{X} \sim N(\mu, \sqrt{\frac{0.5}{n}})$, Then $\frac{\overline{X}\text{-}\mu}{\sqrt{\frac{0.5}{n}}} \sim N(0,1)$

So we have $P(-0.1 < \overline{X} - \mu < 0.1) = P(\frac{-0.1}{\sqrt{\frac{0.5}{n}}} < \frac{\overline{X} - \mu}{\sqrt{\frac{0.5}{n}}} < \frac{0.1}{\sqrt{\frac{0.5}{n}}}) \geq 0.997$

We have the empirical equation that when $\Phi(x) \geq 0.9985$, then $x \geq 2.96$.

So, when $P(\frac{\overline{X} - \mu}{\sqrt{\frac{0.5}{n}}} < \frac{0.1}{\sqrt{\frac{0.5}{n}}}) \geq 0.9985$ which means $\Phi(\frac{0.1}{\sqrt{\frac{0.5}{n}}}) \geq 0.9985$

We can have $\frac{0.1}{\sqrt{\frac{0.5}{n}}} \geq 2.96$, so $n \geq \frac{29.6^2}{2} \approx 438$

n should be at least 438