Assignment 1: Statistics

Maimak Biswas

10 March 2024 18:20

[25 Points – 3 hours]

1) Discrete Distributions – Negative Binomial [5 Points]

In order to get a driving license in country A, a person needs to pass the same driving exam N times. The probability of passing the exam for a 'learner' is P

[Assume Independent attempts]

(i) Define a random variable I_j for success at an j^{th} attempt

(ii) Define a r.v X for number of attempts required to obtain the driving license.

(iii) Find the PMF : $\mathbb{P}(X = k)$

(iv) Find EX : [Hint : use another r.v that defines # trials for 1^{st} success – Geometric – linearity of Exp. and symmetry]

(v) Find $Var(X)$

2) Universality of Uniform – for discrete r.v. [5 points]

Say, $X \sim Bern(P)$

Lets see how we can use uniform distr. to sample from it.

(i) find $F_X(x) = P(X \leq x)$ ← [Instrumental for using Universality]

Now: Recall if : $Y = F_X(x)$

$F_Y(y) = y$ (in case of Continuous : Invertible F)

Because :

$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y)$

$= P(X \leq F_X^{-1}(y)) = y$

[∵ y is invertible]

[if y is invertible]

Else: Define $\tilde{F}_X^{-1}(y) = \min\{x : F_X(x) \geq y\}$

[define inverse as] $0 < y < 1$

and $\begin{cases} \tilde{F}_X^{-1}(0) = -\infty \\ \tilde{F}_X^{-1}(1) = \infty \end{cases}$

(ii) Hence find and plot:

$$P(X \leq \tilde{F}_X^{-1}(y)) = F_X(\tilde{F}_X^{-1}(y))$$

for $X \sim Bern(p)$

(iii) As seen Before you can now sample from

$U \sim Unif(0,1)$

Recall: To, Sample according to distribution $X \sim F_X$

it is essentially : $X = F_X^{-1}(u)$

So here it is : $X = \tilde{F}_X^{-1}(u)$

Argue that this holds for discrete cases like $X \sim Bern(p)$

(iv) Hence write a python code to show it is a valid sampling. (sampling uniform - sample Bernoulli)

3) Joint Distribution, Conditional distribution, Conditional Expectation [5 points]

Consider uniform density in the following region in 2 r.v. X, Y

$$x + y/4 \leq 1 \qquad x \geq 0, \; y \geq 0$$

(i) Find the joint pdf $f_{X,Y}(x,y)$

(i) find the joint pdf $f_{X,Y}(x,y)$

(ii) find conditional distribution $f_{Y|X}(x,y)$

(iii) Hence : $\mathbb{E}[Y/X]$;

(iv) Hence use iterated Expectation to find $E[Y]$

(v) Finally : find $F_X(x)$ [CDF] and use universality of uniform to write a Python script to sample from $F_{X,Y}$.

Hint: sample from F_X } (Universality of uniform)
sample from $F_{Y|X}$ }

4) Transformation of r.vs., MGFs, Bayes' rule [5 points]

(i) Let $Z = \sum_{i=1}^{N} X_i$. X_i's are iid and $N : \Omega \to \mathbb{Z}_+$ is an independent r.v. of X_i's : M_X and M_N are the MGFs of X_i's and N respectively.

S.T. $$M_Z(t) = M_N(\log M_X(t))$$

(ii) $Z \sim N(0,1)$

$$Y_1 = Z^3, \quad Y_2 = Z^4$$

Find pdfs $f_{Y_1}(y)$, $f_{Y_2}(y)$

(iii) $X|N \sim \text{Bin}(N, p)$

$N \sim \text{Poiss}(\lambda)$

(iii) $X|N \sim Bin(N, p)$

$N \sim Poiss(\lambda)$

Find PMF: $P(X = k)$

5) Inequalities, Advanced distributions, Statistical tests [5 points]

(i) Show - Chernoff's inequality,

$$\mathbb{P}(X \geq \epsilon) \leq e^{-\theta\epsilon}\, \mathbb{E}\, e^{\theta X}$$

[Hint : Markov Inequality]

(ii) Let $\chi^2(2) = Z_1^2 + Z_2^2$ [Chi - squared distn]

Find $f_{\chi^2(2)}(y)$

(iii) Say the mean weight for adult women in a country is 58 kg. Now 40 people who eat a particular junk food has an average weight of 66 kg.

(a) Case 1: Say the variance of weight is known to be 5 kg,

(b) Case 2: Now only the sample variance is known for these 40 people - 7 kg.

Find whether this junk food increase weight in individuals. [95 % confidence]

[Define a proper Null hypothesis and use proper tests to accept / discard the null hypothesis (from z / t-test)

You can formulate the problem on paper and use the computer to write