Group 2 (DA02)Theory Assignment(TA04)

I.Concept Questions

1. The experiment is drawing 10 cards without from a standard deck of 52 cards. The cards are of two colors - black (spades and clubs) and red (diamonds and hearts), four suits (spades, clubs, diamonds, hearts), 13 values (2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, King, Ace).
   Random Variable: number of spades drawn
   range- [0,10] (included)

2)
A probability distribution is a mathematical description of a random phenomenon in terms of the probabilities of events. The sample space may be the set of real numbers or a higher-dimensional vector space, or it may be a list of non-numerical values

If you take a sample, each sample having two (or more) observations, from a larger population (or from a probability distribution), the means (and other statistics) of those samples would have a distribution.

This sampling distribution would not be the same distribution as the distribution of the original population. For example, if your original probability distribution was uniform, then the sampling distribution (of the means of the samples) would be approximately

A sampling distribution also describes the set of all possible values for something along with a way to describe how (relatively) likely each of those values is to occur. But the something isn't the value of some property of a randomly selected individual. Instead, it's the value of some statistic which is a function of some collection of randomly selected values from the population.

3)
(a)
a $\in\{...,-2,-1,0,1,2,...\}$,
b $\in\{... ,-2,-1,0,1,2,... \}$, (b>= a)

$n=b-a+1$, total number of values the random variable can take
(b)
$n \in N0$ — number of trials
$p \in [0,1]$ — success probability in each trial
$k \in \{ 0, \ldots, n \}$ — number of successes
(c)
$\lambda > 0$ (real)
$k \in \mathbb{Z}^*$

4)(a)
-infty <a<b<infty ,
$x \in [a,b]$
$f(x) = c$
(b)
$\mu \in R$ — mean (location)
$\sigma 2 > 0$ — variance (squared scale)
$x \in R$

5)
(a) P(X=r) =(nCr)* p^r*(1-p)^(n-r)
    n: total number of trials
    r: number of successes
    p: probability of success in each trial
(b)P(Xi=xi) = n!/(x1! *x2! *...xk!) * p1^x1*p2^x2....pk^xk
    n: total number of trials
    pi: probability that X=Xi
    Xi: categories with success probalities pi
(c) P(X=k) =(Kck)*((N-k)C(n-k)) / (Ncn)
    N: population size
    K: N contains exactly K successes
    k is the successes in n draws,
(d) ℙ(Y1=y1,Y2=y2,…,Yk=yk)= ((m1y1)(m2y2)⋯(mkyk)) /(mn),(y1,y2,…,yk)∈ℕ^k
with ∑(i=1-k)yi=n note: (ab)=aCb
    Yi denote the number of type i objects in the sample
    Di denote the subset of all type i objects and let mi=#(Di)
(e) P(X=k) = λ^k * e^(-λ) / k!
    K: the number of occurrences.
    λ : is the expected number of occurrences.

(c)
z = (X - μ) / σ (X ∈ normal distri.)
or

μ	=	0
σ^2	=	1

6. a. Normal Distribution: f(x)=1/(σ√2π) e^(- 〖(x-μ)〗^2⁄(2σ^2)) -∞<x<∞ Where π=3.14159… and e=2.71828 and μ is the mean and σ is the variance
b. Standard Normal Distribution: f(z)=1/√2π e^(-1/2 z^2) -∞<x<∞ Where π=3.14159… and e=2.71828
c. Chi-squared Distribution: f(x:v)={■(1/(2^(v/2) Γ(v/2)) 〖x^(v⁄2-1) e〗^(-x/2) , x>0 @ @0 Otherwise)┤ Where v is the degrees of freedom of the distribution.
d. Gamma Distribution: f(x:α,β)={■(1/(β^α Γ(α)) x^(α-1) e^(-x/β) x>0@ @0 Otherwise)┤ Where α,β are the shape and rate parameters.

7.
a. Discrete Uniform Distribution: Mean: (a+b)/2 Variance: ((b-a+1)2-1)/12
b. Binomial Distribution: Mean: np Variance: np(1-p) Where n is the number of trails and p is the probability of success.
c. Poisson Distribution: Mean: λ Variance: λ Where λ is the average number of outcomes per unit time
d. Continuous Uniform Distribution: Mean: (a+b)/2 Variance: (b-a)2/12 Where a and b are the starting and ending numbers of the range over which Discrete uniform distribution is applied.
e. Normal Distribution: Mean: μ Variance: σ^2
f. Chi-squared Distribution: Mean: K Variance: 2K Where k is the degree of freedom
g. Gamma Distribution: Mean: α/β Variance: α/β2
h. Weibull Distribution: Mean: λΓ(1+1/k) Variance: λ2 [Γ(1+2/k) – (Γ(1+1/k))2]

8.
a. Normal Distribution: Distribution of heights of people
b. Binomial Distribution: Outcomes of tossing a coin n number of times
c. Poisson’s Distribution: Failure time of a machine in one month
d. Chi-squared Distribution: distribution of mean of a normally distributed population
e. Weibull Distribution: Failure time of a machine component
f. Gamma Distribution: size of insurance claims
g. Hyper geometric Distribution: an urn with two types of marbles. The probability of drawing green marbel without replacement follows hyper geometric distribution.

9. (a)
Differences between Normal Distribution and Standard Normal Distribution is
mean and standard deviation of normal distribution are any real numbers
whereas for standard normal distribution it is 0 and 1 respectively.
9. (b)
Differences between Binomial distribution and Multinomial distribution is
if every event has only two possible outcomes then binomial distribution is used
else if it has many possible outcomes then multinomial distribution is used.
9. (c)
Differences between Binomial distribution and Hyper Geometric distribution is
if the sampling done from the population is without replacement, i.e the events
are dependent on each other then HyperGeometric distribution is used, in contrast
binomial distribution where each event is independent of each other
9.(d)
Differences between Hypergeometric distribution and Multivariate Hypergeometric distribution is
in Hyper geometric distribution there are only two outcomes for any single trial, but multivariate
hyper geometric distribution is considered for experiments in which single trial has more than
two outcomes.

10. (a) 0.1314
10. (b) 0.7372
10. (c) 0.3686

11. mean = 75
sd = 10
P(X > t') ~= 0.1 [~= means "close to"]
=>P(X <= t') ~= 0.9
from the standard normal distribution cumulative function table
we see that t' = 1.28 (approx.)
=>t' = (t-mean)/sd = 1.28
=>t = cut off marks = 87.8

12. A supplier supplies 8 pcs to a retail outlet, which contains 3 of them are defective. If an office makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.
p(no of detectives=0)=5*4/64=5/16
p(no of detectives=1) =3*5/64+ 5*3/64=15/32
p(no of detectives =2) =3*2/64=3/32

13. Classify the following random variables as discrete or continuous:
X: the number of automobile accidents per year in Virginia. - discrete
Y: the length of time to play 18 holes of golf. - continuous
M: the amount of milk produced yearly by a particular cow. -continuous
N: the number of eggs laid each month by a hen. -discrete
P: the number of building permits issued each month in a certain city. discrete
Q: the weight of grain produced per acre. -continuous

14. Determine the value c so that each of the following functions can serve as a probability distribution of the discrete random variable X:
(a) f(x) = c(x2 + 4), for x = 0, 1, 2, 3;
1=c(4+5+8+13),c=1/30
(b) f(x) = c(2Cx)(3C3-x), for x = 0, 1, 2.
c=0.1,it is hypergeometrically distributed C=1/5C3=0.1

15.
a) You want the integral of f(x)dx from x=200 to ∞
=[-10000/(x+100)^2]
= (0) -(-10000/90000)
=1/9

b) You want the integral of f(x) dx from x=80 to 120
=(-10000/220^2) - (-10000/180^2)
=-25/121 +25/81
=1000/9801
=0.1020

16.
(a)P(x<1.2)= integ(xdx) from 0 to 1 + integ(2-x) from 1 to 1.2
=0.5+0.18=0.68
(b) P(0.5<x<1) = integ(xdx) from 0.5 to 1 =0.375

17. Sample space is 1, integ(f(x)dx) from 2 to 5 =1
(a)P(X<4) = 1-integ(f(x)dx) from 4 to 5 =16/27
(b)P(3<=X<4) = 1-integ(f(x)dx) from 4 to 5 -integ(f(x)dx) from 2 to 3 =9/27

18
(a) 1-e^-6(1+6+6^2/2! +6^3/3!
+....6^6/8!)=1-0.0024(7+18+36+54+54*6*(⅖+6/7)+54*27/35)=0.15

(b) e^-6*36/2=0.446

19
(a)C(20,10)*(0.3)^10*(0.7)^10 = 0.0308
C(20,11)*(0.3)^11*(0.7)^9 = 0.0120
C(20,12)*(0.3)^12*(0.7)^8 = 0.004
C(20,13)*(0.3)^13*(0.7)^7 = 0.001
C(20,14)*(0.3)^14*(0.7)^6 = 0.0002
..... <0.0002

0.0308 + 0.012 + 0.004+ 0.001 + 0.0002 + 6*(<0.0002)= 0.0480

(b)
C(20,4)*(0.3)^4*(0.7)^16+C(20,3)*(0.3)^3*(0.7)^17+C(20,2)*(0.3)^2*(0.7)^18+C(20,1)*(0.3)^1*(0.7)^19 +C(20,0)*(0.3)^0*(0.7)^20 =0.23750777

20.Binomial distribution with n = 12 and p = 0.7
Let x be the number of people disapproving.
P(7<=x<=9) = pbinom(9,12,0.7) - pbinom(6,12,0.7) = 0.6293
P(x<=5) = pbinom(5,12,0.7) = 0.0386
P(x>=8) = 1-pbinom(8,12,0.7) = 0.7240
C is most probable.

21.Binomial distribution with n = 9 and p = 0.75
Let x be the number of cars within the state
P(x<4) = pbinom(3,9,0.75) = 0.009994507

22.A total of 8 permutations are possible , only with 2 of them being all same .
HHH/TTT. So, the probability a toss is repeated is ¼.
For less than 3 tosses, the probability is = 1 - (¼)^3 = 63/64 .

23.Assuming the error distribution to be poisson distribution.
(a).mean(μ) = 2
P[x>=4] = 1 - P[e <= 3]
P[x <= 3] = e^-2(1 +2 +2^2/2! +2^3/3!) = 0.85712
P[x >= 4] = 1 - 0.85712346049 = 0.14287

(b).P[0] = e^-2(1) = 0.13533

24.Assuming the hurricane distribution to be poisson with mean = 6.
(a).P[x<=4] = e^-6(1+6 +6^2/2! +6^3/3!) = 0.15120
(b).P[6 <= x <= 8] = e^-6(6^6/6! + 6^7/7! + 6^8/8!) = 0.40155

25. The given observations are 15, 7, 8, 95, 19, 12, 8, 22, and 14.
The mean of the observations is 22.22.
The median of the observations is 14.
The mode of the observations is 8.
The median appears to be the best measure of the centre since by observation, we can see that the mean is is skewed as only one value is greater than it and the rest are all smaller. Similarly, the mode is also skewed since, there is only one value smaller than it and the rest are all larger. In this respect, the median seems to be the most centered of the three tendencies.

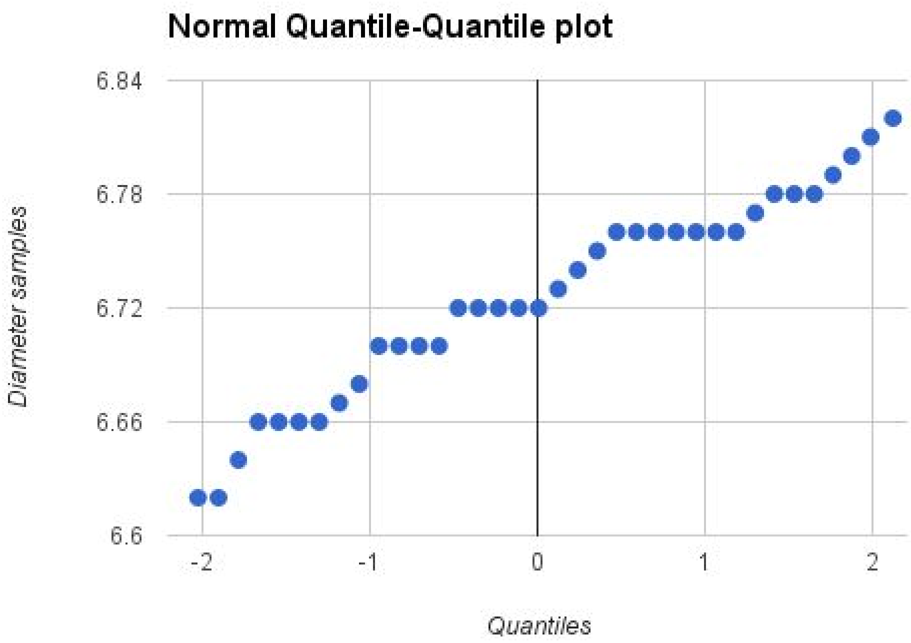
26. Unknown reference 'Exercise 8.3'

27. (a). Mean (μ) = 4 * 0.2 + 5 * 0.4 + 6 * 0.3 + 7 * 0.1 = 5.3
Variance(σ^2) = ($\sum x^2$ * P(X = x)) - μ^2 = (16*0.2 + 25*0.4 + 36*0.3 + 49*0.1) - 5.3^2
= 0.81
(b). Since expectation is independent of X, μ(X) = μ = 5.3
σ^2(X) = (1/n^2)σ^2($\sum$X) = (1/n)*σ^2 = 0.81/36 = 0.0225
(c). P[X < 5.5] = P[(X - μ) / σ < 5.5 - 5.3 / 0.15]
= P[Z < 1.333]
= 0.9082

28. Variance of (n-1)S^2/σ^2 = (n-1)^2/σ^4 * variance of S^2
We know that variance of (n-1)S^2/σ^2 = 2*(n-1)
=> variance of S^2 = 2* (n-1) * σ^4 / (n-1)^2 = 2*σ^4/(n-1)
Therefore, we can see that as n increases i.e. the size of the population increases,
the variance of S^2 decreases.

29. The distribution is a t-distribution with 8 degrees of freedom.
P[t > (Xbar - μ) / (σ/sqrt(n))] = P[t > (24 - 20)/4.1/3]
= P[t > 2.926]
= 0.0095
Therefore, one is not likely to obtain a population of mean 24 and s.d 4.1.
This must imply that either the population is not normal distribution or the values of mean or s.d are incorrect.

30.



II.MCQ's

1.A
2.C
3. D
4. D
5.(A)-W
  (B)-Y
  (C)-Z
  (D)-X
6.D
7.C
8.A
9.B
10.C