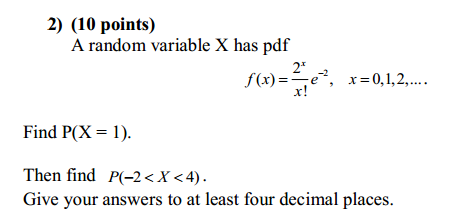


**Solution:** As X is a continuous variable defined with x>1 , we compute P(1<X<4) taking integral values within the range. Thus option b) is the answer



> integrate (function(x) 2\*exp(-2\*(x-1)), lower=1, upper=4)

0.9975212 with absolute error < 1.1e-14



**Solution:**

**P(X = 1)** = f(1) = 21 / 1! \* e-2 = 2e-2

> 2\*exp(-2)

[1] 0.2706706

Thus, P(X = 1) = 0.2706706

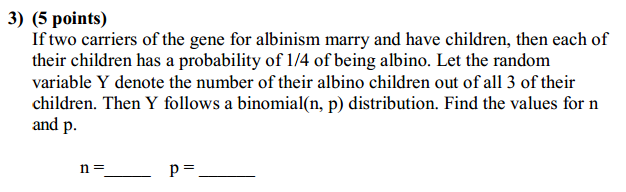
**P(-2 < X < 4)** = f(0) + f(1) + f(2) + f(3) = ∑3x=0 2x /x! \* e-2

> f\_x<- function(x) (2^x/factorial(x))\*exp(-2)

> sum(f\_x(0:3))

[1] 0.8571235

Thus, P(-2 < X < 4) = 0.8571235

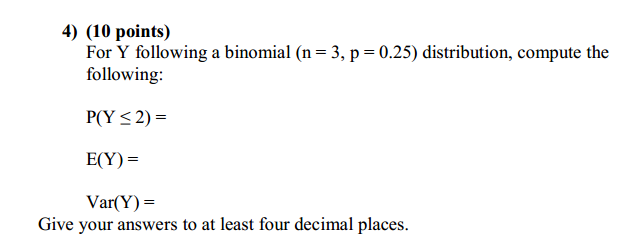


**Solution:**

If n = number of trials i.e. birth of a child

If p = the probability of success i.e. children having albinism

We get the answer as n = 3 and p = ¼ = 0.25



**Solution:**

P(Y ≤ 2) = 0.984375

> set.x<- c(0,1,2)

> sum(dbinom(set.x, size=3, p=0.25))

[1] 0.984375

E(Y) = np for binomial distribution

E(Y) = (3)(0.25) = 0.75

> 3\*0.25

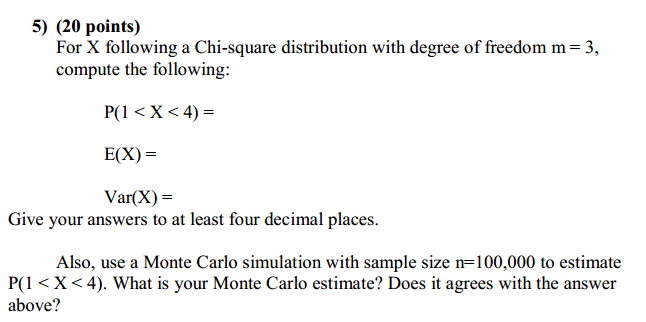
[1] 0.75

Var(Y) = np (1 – p) for binomial distribution

Var(Y) = (3)(0.25)(1 – 0.25) = 0.5625

> 3\*0.25\*(1-0.25)

[1] 0.5625



**Solution:**

P(1 < X < 4) is a continuous variable . it can be computed in two ways .

> # Method 1: cumulative Distribution Function

> pchisq(4,3)-pchisq(1,3)

[1] 0.5397878

> # Method 2: Integrating the pdf

> integrate( function(x) dchisq(x,3), lower=1, upper=4)

0.5397878 with absolute error < 3.1e-14

P(1 < X < 4) = 0.5397878

For Chi-square distribution

E(X) = m = 3

Var(X) = 2m = 6

b) Using monte carlo simulation with size n = 100000 we get

estimate value as 0.53933

> # b) use a Monte Carlo simulation with sample size n=100,000 to estimate

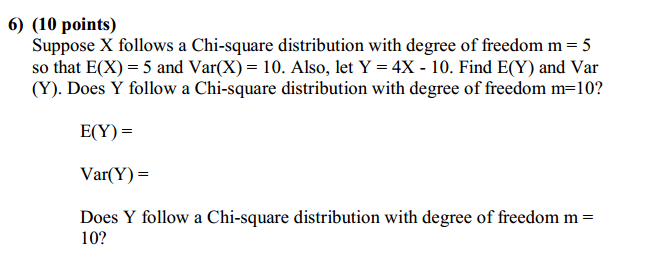
> # P(1 < X < 4).

> X<-rchisq( n=100000, df=3)

> mean((1<X)&(X<4))

[1] 0.53933

**The estimate value we get from monte carlo is approximately same as the value obtained with Chi-square**



**Solution:**

1. Appy the properties of mean and variance of linear transformation of random variables

E(Y) = aE(X) + b

Var(Y) = a2Var(X)

We calculate :

E(Y) = aE(X) + b = (4)(5) – 10 = 20 – 10 = 10

Var(Y) = a2Var(X) = 42(10) = 16(10) = 160

1. Does Y follow a Chi-square distribution with degree of freedom m = 10?

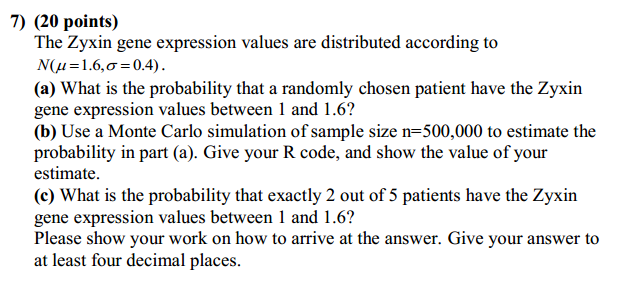
If y followed a Chi-square distribution then,

E(Y) = m = 10

Var(Y) = 2m = (2)(10) = 20

The value of Var(Y) for linear transformation is = 160

Thus, Y **does not** follow Chi-Square distribution with degree of freedom m = 10



**Solution:**

1. Probability for Zyxin gene expression for values between 1 and 1.6 i.e.

P(1 ≤ X ≤ 1.6) can be computed in two methods

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| > integrate (function(x)dnorm(x, mean=1.6, sd=0.4), lower=1, upper=1.6)$value  [1] 0.4331928  > pnorm(1.6, mean=1.6,sd=0.4)- pnorm(1, mean=1.6, sd=0.4)  [1] 0.4331928  Therefore, P(1 ≤ X ≤ 1.6) = 0.4331928   1. Using monte carlo simulation of sample size n=500,000 to estimate the probability   > X<-rchisq( n=100000, df=3)  > mean((1<X)&(X<4))  [1] 0.54136  >  The value of my estimate is 0.54136   1. To find the probability that exactly 2 out of 5 patients have the Zyxin gene expression values between 1 and 1.6,   We can define each patient as a trial and gene expression between 1 and 1.6 as success , thus it can be computed as binomial distribution with  n = 5  p = 0.4331928  for P(X = 2) =   |  |  |  |  |  | | --- | --- | --- | --- | --- | | |  | | --- | | > dbinom(2, size=5, prob=0.4331928)  [1] 0.3417185 | |  | | |  | | --- | |  |   Therefore, P(X = 2) = 0.3417185    **Solution:**  > # Problem 8  >  > # (a) Hand in a R script that calculates the mean and variance of two random  > # variables X~F(m=2,n=5) and Y~F(m=10,n=5) from their density functions.  >  > EX<- integrate(function(x) x\*df(x, df1=2, df2=5), lower=0, upper=Inf)$value  > EX  [1] 1.666667  > EY<- integrate(function(y) y\*df(y, df1=10, df2=5), lower=0, upper=Inf)$value  > EY  [1] 1.666667  > VarX<- integrate(function(x) (x-EX)^2\*df(x, df1=2, df2=5), lower=0, upper=Inf)$value  > VarX  [1] 13.88889  > VarY<- integrate(function(y) (y-EY)^2\*df(y, df1=10, df2=5), lower=0, upper=Inf)$value  > VarY  [1] 7.222222  >  > # (b) Use the formula in Table 3.4.1 to calculate the means and variances  > # directly.  >  > # mean= n/(n-2)  > mean<-5/(5-2)  > mean  [1] 1.666667  > #variance= (2\*n^2\*(m+n-2))/(m\*(n-2)^2\*(n-4))  > variance<- function(m,n) (2\*n^2\*(m+n-2))/(m\*(n-2)^2\*(n-4))  > variance(2,5)  [1] 13.88889  > variance(10,5)  [1] 7.222222 | | |  | | |  | | --- | |  | |   All the outputs are as follows and the answers from part a) agree  With those of part b)   |  | | --- | | > EX  [1] 1.666667  > EY  [1] 1.666667  > VarX  [1] 13.88889  > VarY  [1] 7.222222  > mean  [1] 1.666667  > variance(2,5)  [1] 13.88889  > variance(10,5)  [1] 7.222222 | |  | | |  | | --- | | > | | |
|  |
| |  | | --- | |  | |