Paper Code: BS-110 Paper:
Probability and Statistics for Engineers

List of Practical

1. Fitting of binomial distributions for n and p given.
2. Fitting of binomial distributions, computing mean and variance.
3. Fitting of Poisson distributions for given n and λ.
4. Fitting of negative binomial distribution.
5. Problems related to area property of normal distribution.
6. Fitting of normal distribution when parameters are given/not given.
7. Fitting of gamma distribution.
8. Fitting of exponential distribution.
9. Problems based on measures of dispersion.
10. Problems based on moments, skewness and kurtosis.
11. Karl Pearson correlation coefficient.
12. Correlation coefficient for a bivariate frequency distribution.
13. Lines of regressions.
14. Planes of regressions.
15. Construct box plot of given data
16. Represent the given data using histograms.
17. Performing one way ANOVA and two way ANOVA using block design.
18. Building 2^k factorial design.

Avishisht
Avishisht
00419011721

Topic ______ Date ______

1. Fitting of binomial distribution for n and p given

Example → a coin is tossed 5 times what is the probability of getting 0 heads.

Theory → no. of trials = 5
probability of sucess = 0.5

for binomial distribution

$$P = {}^nC_r \, p^r q^{n-r}.$$

P → binomial probability

x → no. of times for a specific outcome within n trial

p → probability of sucess on a single trial

q → probability of failure " " "

n → no. of trials.

$$P(0) = {}^5C_0 (1/2)^0 (1-1/2)^5$$
$$= 1/25 = 0.03125.$$

Code →

```
num_trials=5
prob_sucess=0.5
prob= dbinom (x=0, num_trials, prob_sucess)
prob
```

```
> trial = 5
> trial = 5
> suc = 0.5
> prob = dbinom(x = 0, trial, suc)
> prob
[1] 0.03125
> 
```

Avishisht
004

Avishisht
004

Topic ____ Date ____

2. Fitting of binomial distribution, computing mean and variances.

Example → consider an exam which consist of 25 MCQ each question has 5 possible answers. This means that probability of answering a question correctly by chance is 0.2

Theory →

n → no. of trials = 5

p → probab. of sucess on each trial = 0.2

no. of observation = 25

mean = n * p (Theoretical)

= 5 × 0.2 = 1

variance = n * p * (1 − p) (Theoretical)

= 5 × 0.2 × 0.8

= 0.8

Code →

```
x = rbinom(25, 5, 0.2)
mean(x)
var(x)
```

~~#~~

```
> x = rbinom(25, 5, 0.2)
> table(x)
x
 0  1  2  3
10  9  4  2
> x
 [1] 0 1 0 0 0 1 0 3 1 1 3 0 2 0 2 0 1 1 1 0 1 2 2 1 0
> mean(x)
[1] 0.92
> var(x)
[1] 0.91
>
```

Topic _______ Date _______

Avishisht
204

Topic _______ Date _______

3- Fitting of poison distribution for given n and λ-

Example→ Calls to a customer-service line came at an average rate of every 5 minutes.
Stimulate the number of calls in 50 such period

Theory→
rpois() function in R is used for generating random numbers from a given poison's distribution
Here n=50 (no. of values of return)
$\lambda=6$
The poison distribution has a density:

$$p(x) = \frac{\lambda^{x} e^{-\lambda}}{x!} \quad \text{for } x=0,1,2\ldots$$

Code

```
num_values =50
lambda=6
observations=rpois (num_values,
                    lambda)
table(rpois(50,6))
```

```
> num_values = 50
> lambda = 6
> observations = rpois(num_values, lambda)
> table(rpois(50, 6))

 1  3  4  5  6  7  8  9 10 11 12
 2  7  8  4  8  6  8  2  3  1  1
> 
```

Avishisht
004

Topic ____________ Date ____

4. Fitting of negative binomial distribution

~~[illegible]~~

Theory → negative binomial distribution describes the number of failed Bernoulli trials that occur before r successes. The probability mass function for such a random variable is.

$$f(x) = P(X=x) = \binom{x+r-1}{r-1} p^r q^x$$

p → probability of sucess on an individual trial and $q = 1-p$ is probability of failure.

Code →

```
num_obs = 100
size = 6
prob_sucess = 0.2
boxes <- rnbinom
table(boxes)
```

```
> num_obs = 100
> size = 6
> prob_success = 0.2
> boxes <- rnbinom (num_obs, size, prob_success)
> table(boxes)
boxes
 5  6  7  8  9 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 31 33 34 35 36 40 41 42 44 46 49
 1  2  1  2  2  4  2  3  4  4  5  4  2  5  4  1  7  7  5  1  2  4  4  1  2  2  3  1  4  1  1  1  1  1  3
51 57 71
 1  1  1
> 
```

Avishisht
004

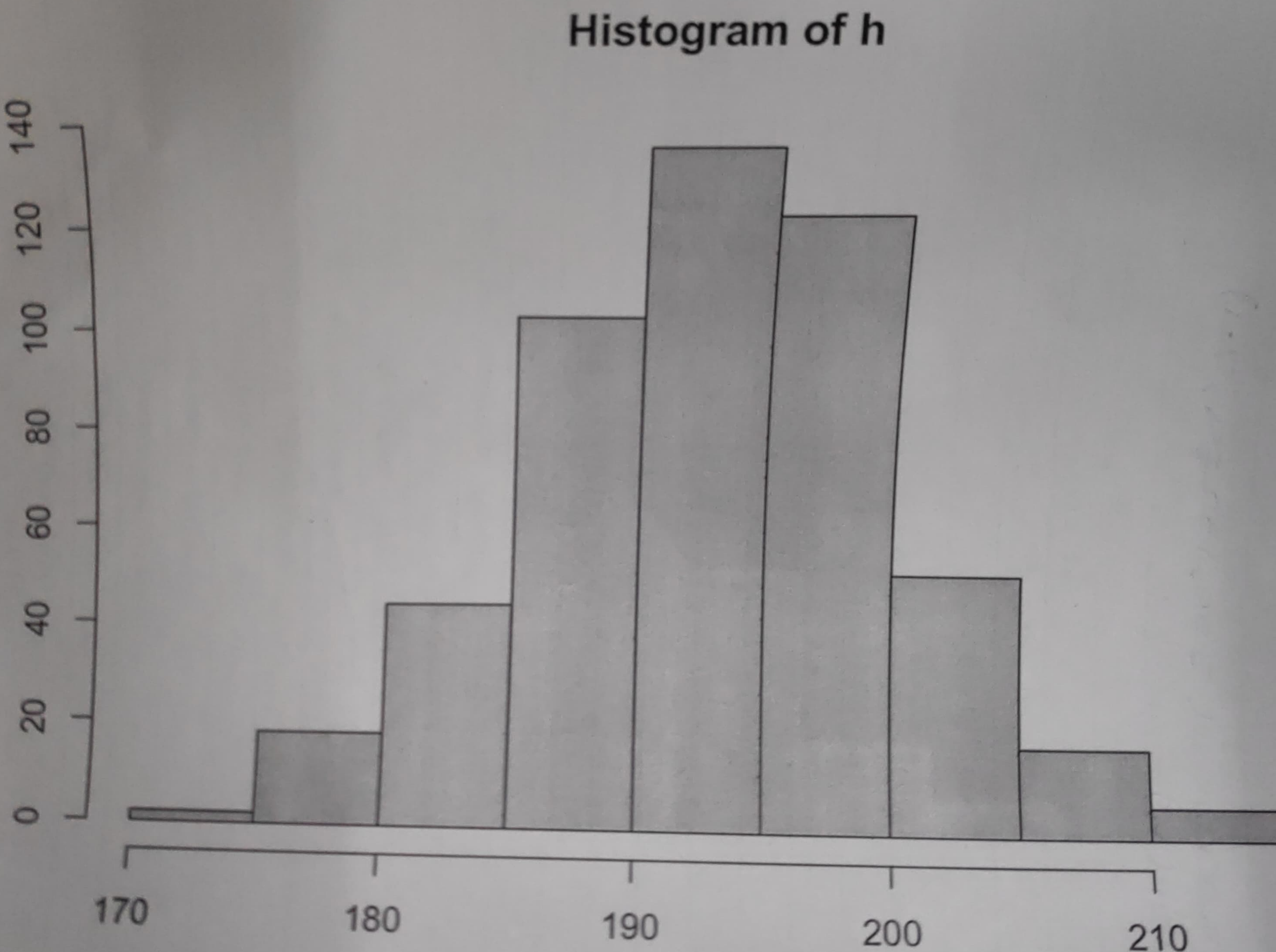
Topic Date

Avishisht
004

DELTA Pg No.
Date / /

Q5 Problems related to area property of normal distribution

```
h=rnorm(500,192.9,7.1)
hist(h)
```


Histogram of h
0
20
40
60
80
100
120
140
170
180
190
200
210

Avishisht
004

Topic_______________ Date_______

6- Fitting of normal distributions when parameters are given / not given?

Theory → mean = 192.9 mm
sd = 7.1 mm
To calculate the probability, we will pnorm function which takes 4 arguments

Code →

```
prob_less200 = pnorm(200, 192.9, 7.1)
prob_less200
prob_greater200 = 1 - prob_less200
prob_greater200
```

7. Fitting of gama distribution

Theory → Gamma distribution is completely specified by two parameter, θ, representing mean waiting time between events and α, total number of occurance needed.
In this case, we will use pgamma function which takes 3 arguements = x, alpha & lambda
here x = 60 (1 hour)
alpha = 25
lambda = 1/3

Code:

```
x=60
alpha=25
lambda=1/3
prob= 1-pgamma(x,alpha, lambda)
prob.
```

Q5

```
> prob_less200 = pnorm(260, 192.9, 7.1)
> prob_less200
[1] 1
> prob_greater200 = 1 - prob_less200
> prob_greater200
[1] 0
>
```

~~Q6~~. Q7

```
> x = 60
> alpha = 25
> Lambda = 1/3
> prob = 1 - pgamma(x, alpha, Lambda)
> prob
[1] 0.8432274
> |
```

Topic ______ Date ______

8. Fitting of exponential distribution

Theory → Exponential distribution models the waiting time b/w identical and independent randomly occuring events like calls to a pizza. Exponential distribution is specified by mean number of occurrences in a interval. In this case, we will use pexp (x, lambda) which returns probability that the waiting time b/w two occurance is no more than x.

Code

```
x=3
lambda = 6/5
prob = 1 - pexp (x, lambda)
prob.
```

```
> x = 3
> lambda = 6/5
> prob = 1 - pexp(x, lambda)
> prob
[1] 0.02732372
> 
```

Topic Date

Q9 Problems based on measures of dispersion.

Theory:

mean $\mu = \frac{5+10+15+20+25}{5} = 15$

variance, $\sigma^2 = \frac{100+25+0+25+100}{5-1} = 62.5$

standard deviation $\sigma = \sqrt{62.5} = 7.905$

Range = 5-25

Code:

```
x <- c(5,10,15,20,25)
print(sd(x))
print(var(x))
print(range(x))
print( mean(x))
```

```
> x <- c(5, 10, 15, 20, 25)
> print(sd(x))
[1] 7.905694
> print(var(x))
[1] 62.5
> print(range(x))
[1]  5 25
> print(mean(x))
[1] 15
> |
```

Avishisht

Experiment Name / No.:

Camlin Page No.

Date / /

Q10 Problem based on moment, skewness and kurtosis

Code:

```
#q(10)(1)
x <- c(1,3,4,5,6,7)
library(moments)
kurtosis(x)
#q(10)(2)
x <- c(1,3,4,5,6,7)
library(moments)
print(skewness(x))
```

```
> #q10(1)
> x<-c(1,3,4,5,6,7)
> library(moments)
> kurtosis(x)
[1] 2.04
> #q10(2)
> x<-c(1,3,4,5,6,7)
> library(moments)
> print(skewness(x))
[1] -0.3380617
> |
```

Avishisht
004

Topic Date

Q11. Karl Pearson correlation Coefficient:

- Karl Pearson correlation: The correlatn is an extensively used methamatical method in which the numerical representation is applied to measure the level of relation b/w linearly related variable represent by 'r'

- Cor()
- By using the function we can calculate
- Syntax:
- Cor (x, y, method="Pearson")
- Example: Calculate Karl pearson coefficient of correlation between the following series

HUSBAND'S AGE	21	22	23	24	25	26	27
WIFE'S AGE	16	15	17	18	19	20	21

- solving mathematically:

Avishisht

Avishisht 004

Topic (Q11 continued) Date

X	$x = X - A$	x^2	Y	$y = Y - A$	y^2	xy
21	-3	9	16	-2	4	6
22	-2	4	15	-3	9	6
23	-1	1	17	-1	1	1
24	0	0	18	0	0	0
25	1	1	19	1	1	1
26	2	4	20	2	4	4
27	3	9	21	3	9	9
$\Sigma X = 168$		$\Sigma x^2 = 28$	$\Sigma Y = 126$		$\Sigma y^2 = 28$	$\Sigma xy = 27$

$\therefore \bar{X} = 168/7 = 24$

$\bar{Y} = 126/7 = 18$

Coefficient of correlation $= \dfrac{\Sigma xy}{\sqrt{\Sigma x^2 \times \Sigma y^2}}$

$\therefore$ correlation $= \dfrac{27}{\sqrt{28 \times 28}}$

$\therefore$ Correlation $= 27/28 = 0.96$

- CODE:

```
x <- c(21,22,23,24,25,26,27)
y <- c(16,15,17,18,19,20,21)
cor(x, y, method='pearson')
```

Console Terminal Jobs

R 4.2.0 · ~/

```
> #Creating data set
> x<-c(21,22,23,24,25,26,27)
>
> y<-c(16,15,17,18,19,20,21)
>
> #SYNTAX
> cor(x,y,method="pearson")
[1] 0.9642857
> |
```

Q12 Correlation coefficient for a bivariate frequency distribution

Theory:-

$$r = \frac{n \sum fuv - \sum fu - \sum fv}{\sqrt{n \sum fu^2 - (\sum fu)^2}}$$

$u = \frac{X - A}{h}$ here A will be assumed mean
h → factor for step deviation

$v = \frac{Y - B}{k}$, here B will be assumed mean
k → factor for step deviation.

Code:

```
x <- c(1,2,3,4)
y <- c(5,6,7,8)
result = cor(x,y)
print (result)
```

```
> x <- c(1,2,3,4)
> y <- c(5,6,7,8)
> result = cor(x,y)
> print(result)
[1] 1
> |
```

Avishisht
004

Topic ______ Date ______

13. Lines of regression.

Example → The sales of a company (in million dollars) for each year are shown in table below

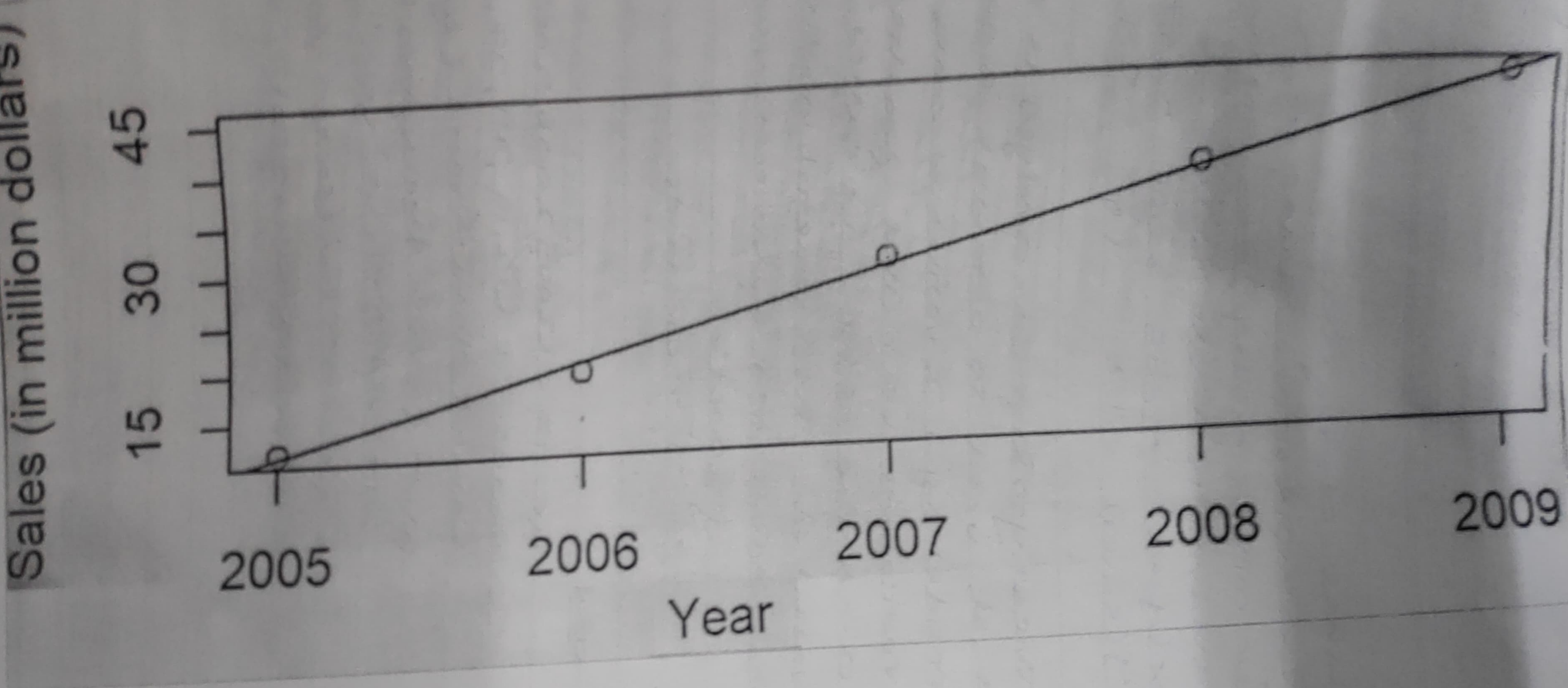
x (year)	2005	2006	2007	2008	2009
y (sales)	12	19	29	37	45

Theory → Regression analysis is statistical tool used to establish a relationship model b/w 2 variables one of these variables is called predictor variable and the other is called response variable. The general mathmatical equation for a linear regression is:-

$$y = \underbrace{ax+b}_{\text{constants}}$$

Code →

```
year = c(2005,2006,2007,2008,2009)
sales = c(12,19,29,37,45)
table = data.frame(year, sales)
plot(table$year, table$sales,
 col="red", xlab='year', ylab='sales')
regression = lm(sales~year, data=table)
abline(regression, col='blue')
```


Sales (in million dollars)
15
30
45
2005
2006
2007
2008
2009
Year

Avishisht
004

Topic ______ Date ______

Q14 Planes of regression

Theory:

$y = a + \beta_1 x_1 + \beta_2 x_2 + \dots$

Here $\beta_i = b\frac{S_x}{S_y}$, β_i is beta coefficient, b is standard regression coefficient, S_x is sample deviation for x and S_y is sample deviation for y.

Code

```
x1 <- c(1,2,3,4)
x2 <- c(5,6,7,8)
y <- c(10,20,30,40)
dataset= cbind.data.frame(x1,x2,y)
ls <- function(dataset, par)
{with(dataset,sum(y-par[1]-para[2]*x1
        -par[3]*x2)^2)}
result <- optim(par=c(0,0,0),ls,data
                                  =dataset)
coef <- resut $ par
print ((coef)
```

Q15 Construct box plot of given data

Code:

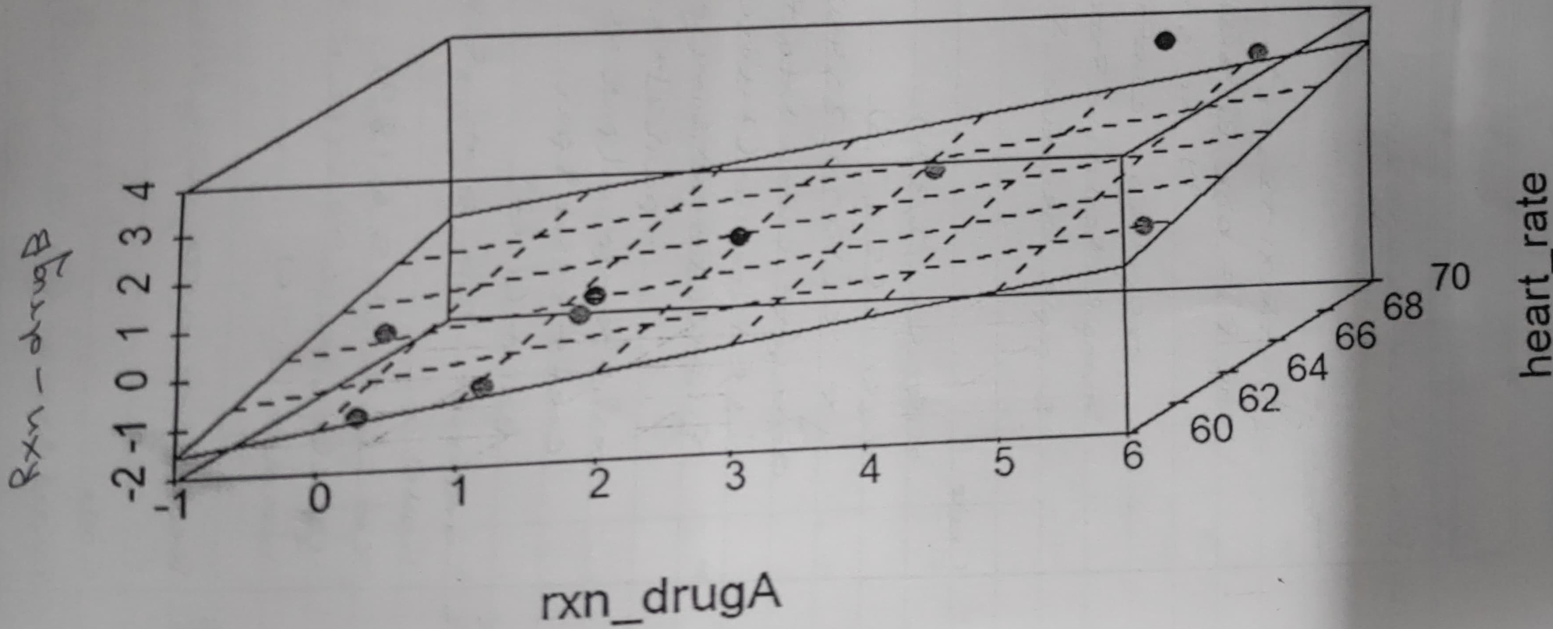
```
x <- c(1,2,3,4,5,6,7,8,9)
y <- c(10,20,30,40,50,60,70,80,90)
boxplot (y~x)
```

Q16 Represent data using histogram

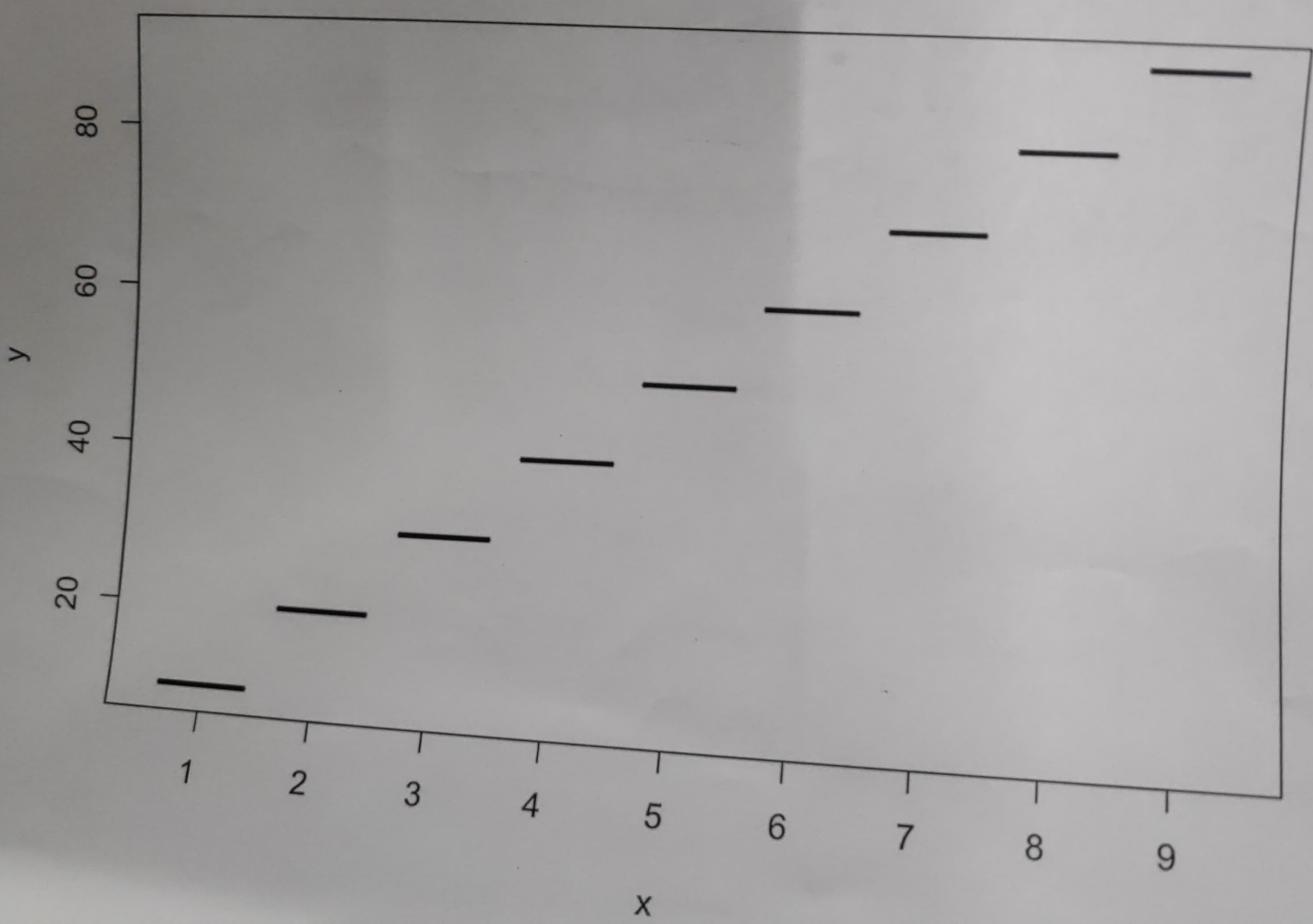
Code:

```
hist (dbinom(0:10,10,0.2))
```

Regression Plane



Q15



Q16

Histogram of dbinom(0:10, 10, 0.2)