Mrinal Chaudhari- Applied statistic- 614

Homework 2 - Normal Distribution and Probability Plots

1. (Show Work) Assume that X is normally distributed with mean 20 and standard deviation 2.
Determine the following:
a. P(X<24)
b. P(X>18)
c. P(18 < X < 22)
d. Now determine x such that $P(X > x)$ = 0.9772

Solution:

a: P(X<24):
μ=20 **σ**=2

Z= $x - \mu/$ = $24 - 20/2$ = 2
P(X<24)=P(Z<2)= 0.97725

b: P(X>18):
μ=20 **σ**=2

Z=Z= $x - \mu/$ = $18 - 20/2$ = -1
P(X>18)=1-P(X<18)=P(Z>-1)= 1- 0.15866= 0.84134

c: P(18 < X < 22):
μ=20 **σ**=2

Z= $18 - 22/2 =- 2$

P(18 < X < 22)= P(X<22)-P(X<18)= 0.84134- 0.15866=0.68268

d. Now determine x such that $P(X > x)$ = 0.9772:

P(X>x)=0.9772
Z= $x - \mu/$ =$x - 20/2$
P(Z>z)=z=2=$x - 20/2$
2=$x - 20/2$
x-20=4
x=24

2: (Show Work) The length of an injected-molded plastic case that holds tape in normally distributed with mean length 90.2 mm and a standard deviation of 0.1 mm.
e. What is the probability that a part is shorter than 90.3 mm and longer than 89.7 mm?

f. What should the process mean be to obtain the greatest number of parts between 89.7 and 90.3 mm?
g. If parts that are not between 89.7 and 90.3 mm are scrapped, what is the yield of the process mean that you chose in part b

Solution:

e. What is the probability that a part is shorter than 90.3 mm and longer than 89.7 mm?

To find P(89.7<X<90.3)=

μ=90.2 σ=0.1

Z= $89.7 - 90.3/0.1 =- 6$

P(89.7 < X < 90.3)= P(X<90.3)-P(X<89.7)= 0.84134-0=0.84134

f. What should the process mean be to obtain the greatest number of parts between 89.7 and 90.3 mm?

P(89.7<X<90.3)= P(X<90.3)-P(X<89.7)= 0.8413
Greatest number of parts between 89.7 and 90.3= 89.7+90.3/2= 90
Process mean is 90

g. If parts that are not between 89.7 and 90.3 mm are scrapped, what is the yield of the process mean that you chose in part b:

μ= $89.7 + 90.3/2$ =90

P(89.7≤X≤90.3)=P(90-3*(0.1) ≤X≤ 90+3*(0.1))

Equating with empirical formula:'

P(89.7≤X≤90.3)=P((μ-3σ) ≤ X ≤ (μ+3σ))

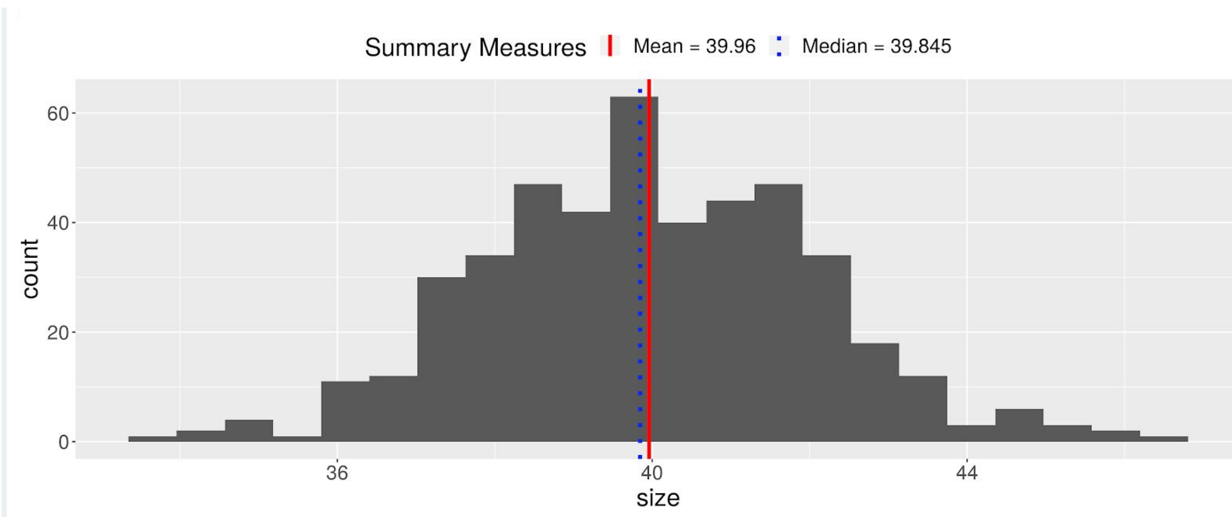
Based on empirical formula,
99.7 % of the data falls in $\mu \pm 3\sigma$
Hence the yield would be **99.7%**

3. DASL displays a histogram of chest sizes of 457 men in the Scottish army in the early 1800s. See

the file chest.txt for the data.
a. Use the data given to create a histogram. Based on the plot, does the data seem follow a normal distribution?
b. Now use a probability plot to decide if the data follows a normal distribution
c. Compute the mean and standard deviation of the data
d. Using your estimates, and assuming that the data does follow a normal distribution, what percent of men do you expect to have chests bigger than 43 inches? What percent actually do?
e. Based on your answers to part d, do you believe that the data comes from a normal distribution? Why or why not?

Solution:

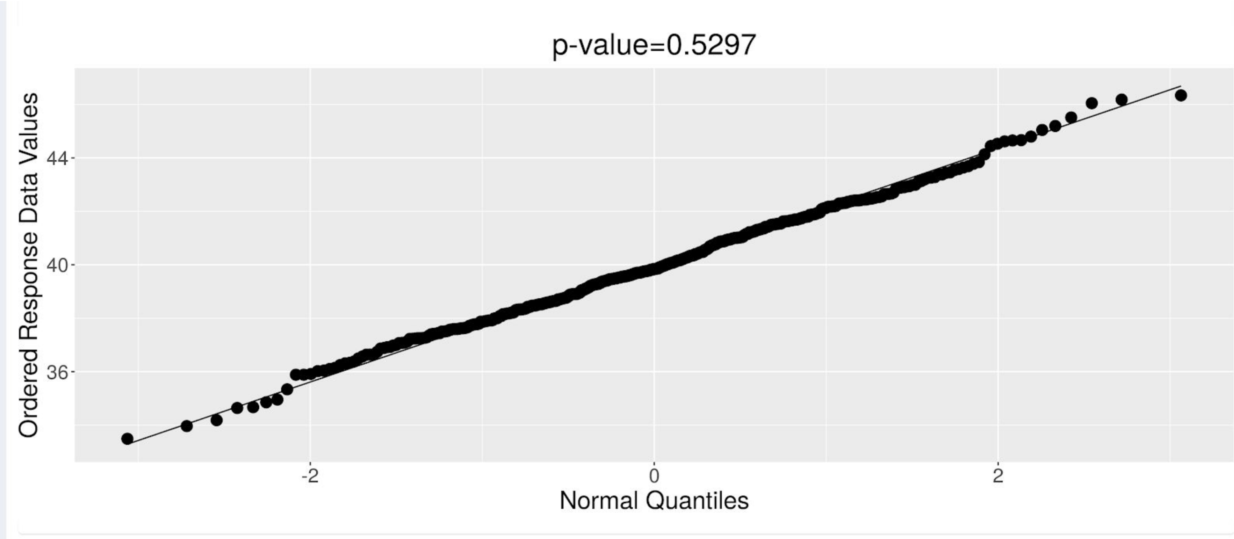
a. Use the data given to create a histogram. Based on the plot, does the data seem follow a normal distribution?



According to the graph, Mean and Median values are quite close to each other hence we can say that the graph follows normal distribution.

b. Now use a probability plot to decide if the data follows a normal distribution

Mrinal Chaudhari- Applied statistic- 614



From the above graph almost all point are lying on the same line i.e it is linear and hence it follows normal distribution

c. Compute the mean and standard deviation of the data

Values of Mean and Standard deviation:

mean=39.96
SD= 2.107104313

d. Using your estimates, and assuming that the data does follow a normal distribution, what percent of men do you expect to have chests bigger than 43 inches? What percent actually do?

Solution:

To check the percentage of men having chest size bigger than 43 inches if the data follows normal distribution:

Actual probability :

That means we need to find: $P(X \geq 43)$
We can write it as: $P(X \geq 43) = 1 - P(X \leq 43)$
So for finding $P(X \leq 43)$:
Standard deviation σ: 2.107
Calculating $Z = (43-39.96)/2.107$

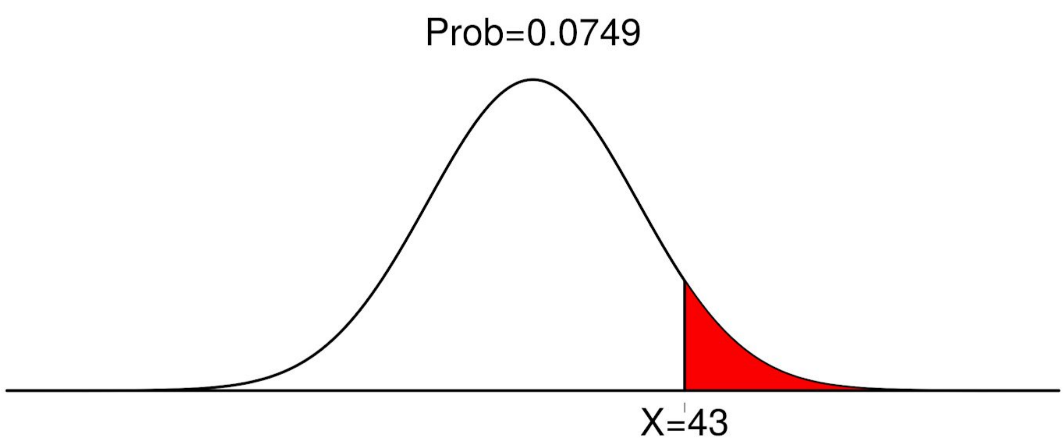
= 3.04/2.107 = 1.44
$P(X \le 43) = 0.9251$
Now $P(X \ge 43) = 1 - P(X \le 43)$
$= 1 - 0.9251$
$= 0.074$
This shows that 7.4% of people have chest size bigger than 43 inches.
Even the bell curve below suggests that the distribution to have a percent of **7.4%** people having chest size greater than 43 inches.

Expected probability:
According to data given in excel file, only **29 people** have chest size greater than 43 inches.
Hence,
$P(X \ge 43) = 29/457 * 100 = 0.06347*100 = 6.3\%$
Actual probability of people having chest bigger than 43 inches is **6.3%**

The probability becomes:

$$P(X \geq 43) = 0.0749$$

We can visualize this with the curve below:



e. Based on your answers to part d, do you believe that the data comes from a normal distribution? Why or why not?

Normal distribution is symmetric about mean hence half data will be less than mean and half data will be greater than mean.
From the above answer (d) it shows that 7.4% of people have chest size greater than 43 inches.
According to the bell curve also we can notice the same probability.
The data has a bell curve distribution and the mean and median are approximately equal to each other. Based on this information, We can say that data comes from normal distribution.