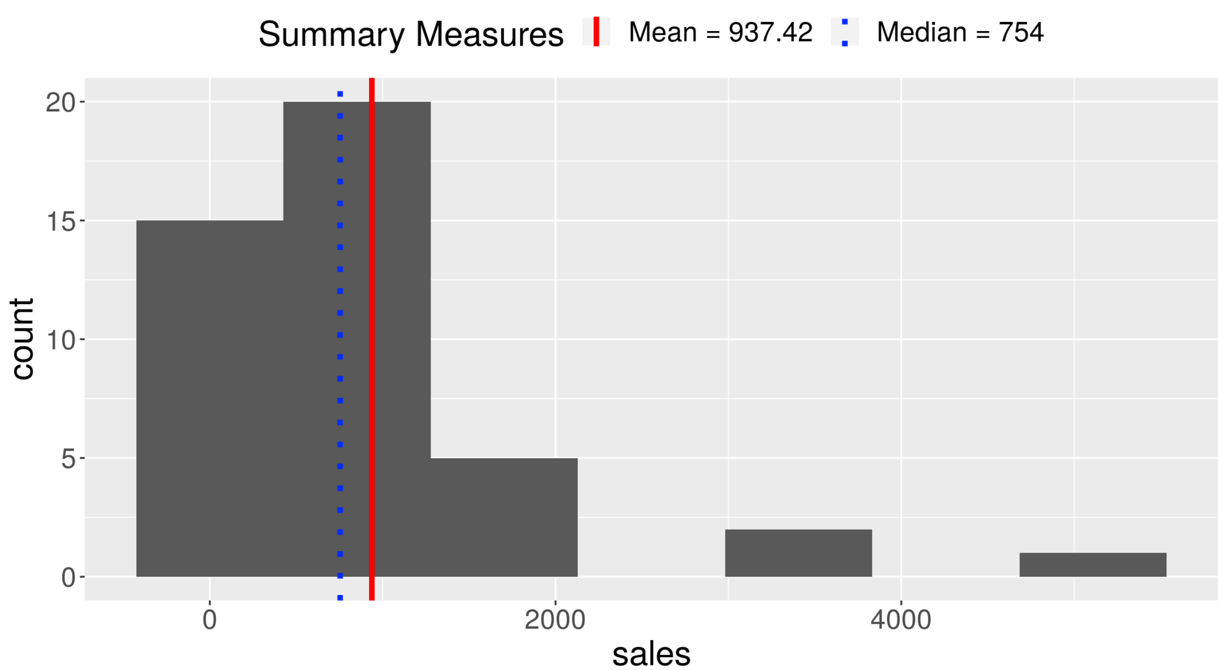
0307-614-02 Term Spring '21 Prof. Parody
Name- Mrinal Chaudhari

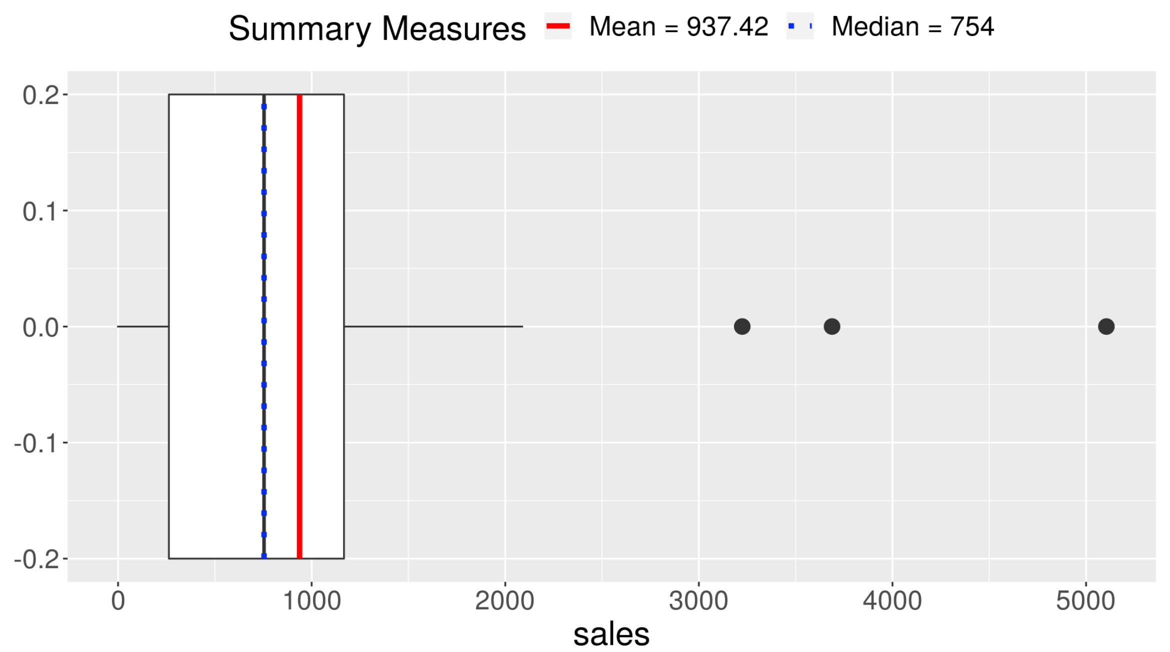
1: You have been given the task to analyze the number of monthly heating equipment orders. The data for roughly 4 years' worth of orders. Use the data to create a preliminary analysis of the data and draw conclusions. This should be solely descriptive in nature and should give a good indication of what is going on with the data. You should make a recommendation based off of the analysis. The data can be found in mycourses in a file called heating.txt.

Solution:

Histogram:



Box Plot:



Mean= 937.42
Median= 754
Variance= 1045206
Standard Deviation= 1022.35
Sample size(n) = 43
Q1=227
Q3=1176
IQR=949
Min=-5
Max= 5105
Range=5110
Lower fence= -1196.5
Upper fence= 2299.5

From the graph and data we can say that as winter comes the sales goes up for 2 consecutive years and later it falls down for the third year.

2: An important characteristic for a reactor is substrate concentration (mg/cm3). Assume that the concentration has a mean = 0.3 and σ2=0.0025. An alarm will go off if the concentration falls outside of the acceptable range which is 0.25 and 0.35. In addition, if

the concentration falls below the lower limit, a report will need to be written and submitted. (please show work for all parts except a)
a) Use the attached probability plot to explain why we can assume that the data comes from a normal distribution?
b) Explain why we have the best situation right now in terms of the acceptable range
c) What is the probability that an alarm will not go off?
d) What is the probability that an alarm will go off but a report will NOT need to be written?
e) What is the probability that both an alarm will go off and that a report will need to be written?

Solution:

Mean=0.3
$\sigma 2$=0.0025

a) Use the attached probability plot to explain why we can assume that the data comes from a normal distribution?

From the probability graph we can say that point lies in the same line which follows a 45 degree normal distribution and there are no outliers present in the graph.

b) Explain why we have the best situation right now in terms of the acceptable range

The acceptable range for concentration is 0.25 and 0.35, if we observe mean i.e 0.3 lies exactly between the given range. Also the probability shown is 0.2765 and variance is 0.0025 which is negligible. Hence we can say that the current situation is best in terms of acceptable range.

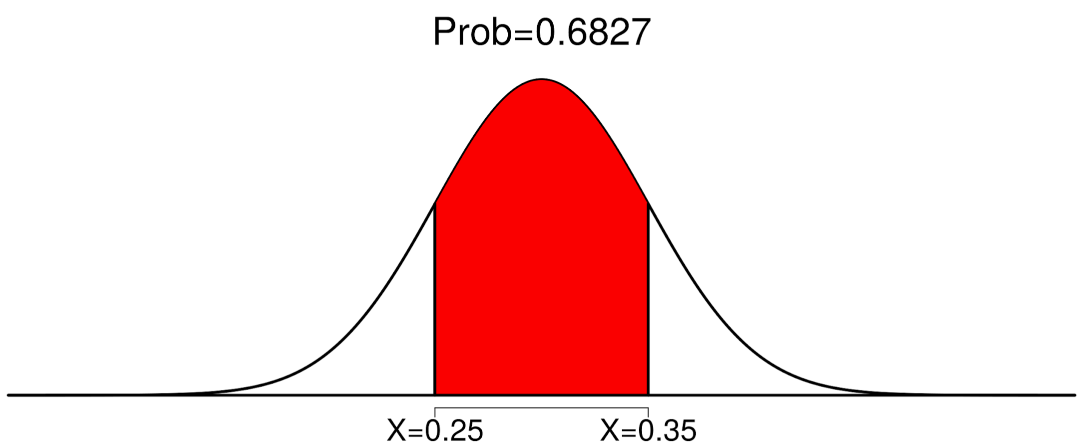
c) What is the probability that an alarm will not go off?
Solution:

Mean=0.3
$\sigma 2$=0.0025
SD= $\sqrt{\sigma 2}=\sqrt{0.0025}$=0.05
P(0.25≤X≤0.35)
Z= x − μ/σ
Z= 0.25-0.35/0.05=-2

P(0.25≤X≤0.35)=P(X≤0.35)-P(X≤0.25)
P(0.25≤X≤0.35)=0.8413-0.1587
P(0.25≤X≤0.35)=0.6827

0.6827 is the probability that an alarm will not go off

We can visualize the curve below:



d) What is the probability that an alarm will go off but a report will NOT need to be written?

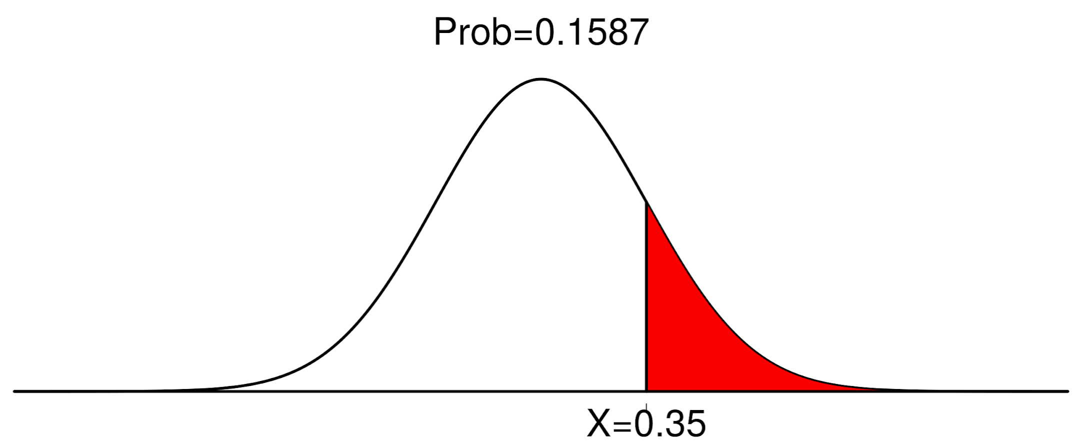
Mean=0.3
σ2=0.0025
SD= $\sqrt{\sigma 2}$=$\sqrt{0.0025}$=0.05
P(X≥0.35)

Z= x − μ/σ
Z=0.35-0.3/0.05
Z= 1

P(X≥0.35)=1-P(X≤0.35)
P(X≥0.35)=1-0.8413
P(X≥0.35)= 0.1587

0.1587 is the probability that an alarm will go off but a report will NOT need to be written

We can visualize this with the curve below:



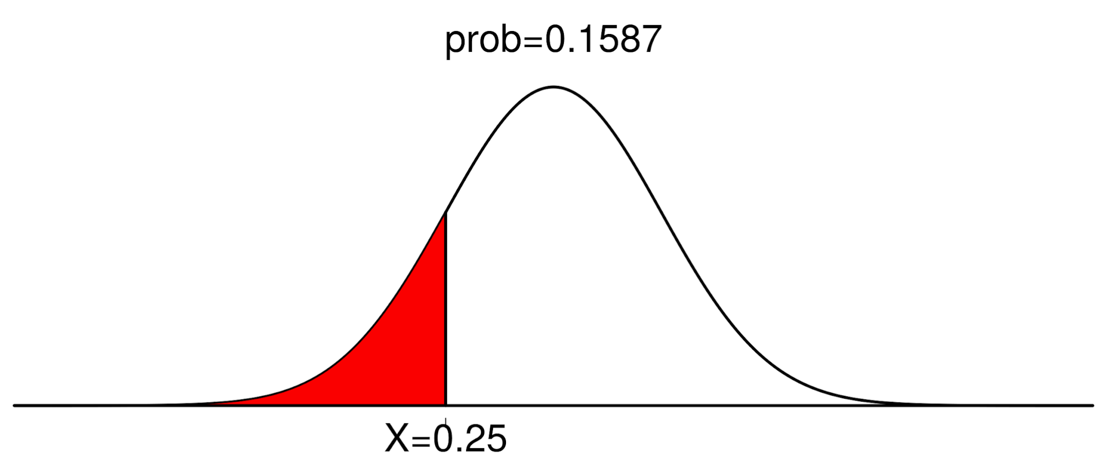
e) What is the probability that both an alarm will go off and that a report will need to be written?

P(X<0.25)=
Z= x − μ/σ
Z=0.25-0.3/0.05
Z= -1

P(X<0.25)=0.1587

0.1587 is the probability that both an alarm will go off and that a report will need to be written

We can visualize this with curve below:



3: Representatives from a certain number of universities are travelling to a convention. There are 4 possible ways to get there, air, bus, car or train. The following study was run on representatives from NY. For the 15 randomly selected delegates from the state of NY, the table below yields the actual probabilities and the counts. (please show work for all this problem)

Mode	**Count**	**Prob**
Air	**3**	**0.4**
Bus	**3**	**0.2**
Car	**5**	**0.3**
Train	**4**	**0.1**

a. What is the actual probability that someone will arrive by land only
b. What distribution does the number of delegates arriving by land follow (including the parameter values)? Explain your answer
c. What is the average number of delegates that will arrive by land only
d. Determine the probability that at least 10 delegates arrive by land only
e. Determine the probability that exactly 5 delegates arrive by air

Solutions:

a. What is the actual probability that someone will arrive by land only

P(Land)=P(Bus)+P(Car)+P(Train)
P(Land)=0.2+0.3+0.1
P(Land)=0.6

b. What distribution does the number of delegates arriving by land follow (including the parameter values)? Explain your answer

It is a binomial distribution.
The success is delegates coming by land and failure is delegates not coming by land.
Also, All trials are independent of each other.
Sample size is 15 i.e the total number of delegates.

c. What is the average number of delegates that will arrive by land only

$n=15$ and $\pi=10/15=0.66$
$\mu=n\pi$
$\mu=15*0.6$
$\mu=9$
The average number of delegates that will arrive by land only is 9

d. Determine the probability that at least 10 delegates arrive by land only:

n=15, π=10/15=0.6
b(x;n,p) =$\binom{n}{x}$ π^{x} $(1 - \pi)^{n-x}$ x=0,1,2….n
P(X≥10)=P(X=10)+P(X=11)+P(X=12)+P(X=13)+P(=14)+P(X=15)
P(X≥10)= $\binom{15}{10} 0.6^{10} (1 - 0.6)^{15-10}$ + $\binom{15}{11} 0.6^{11} (1 - 0.6)^{15-11}$ +$\binom{15}{12} 0.6^{12} (1 - 0.6)^{15-12}$+
$\binom{15}{13} 0.6^{13} (1 - 0.6)^{15-13}$+ $\binom{15}{14} 0.6^{14} (1 - 0.6)^{15-14}$+ $\binom{15}{15} 0.6^{15} (1 - 0.6)^{15-15}$
P(X≥10)=0.1859+0.1268+0.0634+0.0219+0.0047+0.0005
P(X≥10)=0.4032

The probability that at least 10 delegates arrive by land is 0.4032

e. Determine the probability that exactly 5 delegates arrive by air

n=15 and π=0.4(Probability of delegates arriving by air)--given
μ=nπ
μ=15*0.4
μ=6

Using Binomial pmf to find the probabilities
The probability mass function(pmf) for the binomial distribution is:
P(X=X0)=$\binom{n}{x0}$ π^{x0} $(1 - \pi)^{n-x0}$
It becomes,
P(X=5)= $\binom{15}{5} 0.4^{5} (1 - 0.6)^{15-5}$
P(X=5)=0.1859

The probability that exactly 5 delegates arrive by air is 0.1859

4. You are an engineer testing the thickness of eyeglasses for a supervisor. A sample of 10 lenses yields a mean thickness of 2.95 mm and standard deviation 0.34 mm. The engineer would like the true average thickness of such lenses to be less than 3.20 mm. Assume that the underlying distribution is normal and using a type I error rate of 10%:
a) What distribution will you use for the critical point and why
b) Would you create a confidence interval, upper bound or lower bound? (Explain)
c) Find the confidence interval or bound that you choose in part b and interpret it (show work)
d) Give the real world answer

Solution:

Here Total sample n=10
mean=2.95
SD=0.34
SD= $\sqrt{variance}$
Variance= σ^2/n
variance= 0.1156/10
Variance= 0.01156

a) What distribution will you use for the critical point and why

Normal distribution can be used for critical points because μ and σ determine the shape of the distribution. Critical values can tell about the probability that any particular value will have a distribution.

b) Would you create a confidence interval, upper bound or lower bound? (Explain)

n=10
$\overline{x}$=2.95
σ=0.01156
$Z_{0.05}$ = 1.64
Error rate= 10%

(1-α)100% [(1-10)100% i.e 90%] upper confidence bound for μ is given by:

We can put values,

$\mu < \overline{x} + Z\ \alpha * \sigma/ \sqrt{n}$
$\mu < \overline{x} + Z_{0.05} * \sigma/ \sqrt{n}$
$\mu < 2.95 + 1.645 * 0.01156/\sqrt{10}$
$\mu < 2.95 + 1.645 * 0.01156/3.16$
$\mu < 2.95 + 1.645 * 0.00365$
$\mu < 2.95 + 0.00601$
$\mu < 2.95601$

Error rate=10%
(1-α)100% [(1-10)100% i.e 90%] lower confidence bound for μ is given by:

$\mu > \overline{x} - Z_{0.05} * \sigma/ \sqrt{n}$
$\mu > 2.95 - 1.645 * 0.01156/\sqrt{10}$
$\mu > 2.95 - 1.645 * 0.01156/3.16$
$\mu > 2.95 - 1.645 * 0.00365$
$\mu > 2.95 - 0.006004$
$\mu > 2.9439$

From the above values, we can choose the upper bound to find confidence as the engineer would like the true average thickness to be less than 3.20 mm. Hence choosing upper bound may be the correct choice in this case.

c) Find the confidence interval or bound that you choose in part b and interpret it (show work)

n=10
$\bar{x}$=2.95
σ=0.01156

Data is assumed to be normally distributed
And the target value is less than 3.20 mm.
$\mu \in \bar{x} \pm Z\ \alpha/2 * \sigma/\sqrt{n}$

Error rate is 10%
To calculate 90% confidence interval, Z value is 1.645
$\mu \in 2.95 \pm 1.645 * 0.01156/\sqrt{10}$
$\mu \in 2.95 \pm 1.645 * 0.01156/3.16$
$\mu \in 2.95 \pm 1.645 * 0.00365$
$\mu \in 2.95 \pm 0.006004$
$\mu \in (2.956004, 2.9439)$

After calculation of Confidence Interval, We can say that the μ value lies between 2.956004 and 2.9439. we are 90% confident that the engineer will like the average thickness to be less than 3.20 mm .

d) Give the real world answer

The engineer would like the true average thickness of lenses to be less than 3.20 mm. Error rate is 10% hence we calculated 90% confidence interval and upper bound. We assumed that data is normally distributed and values lie between 2.956004 and 2.9439. Our target value i.e 3.20 falls in that interval.