Mid-term exams will be on 23/11/2021, 18:00

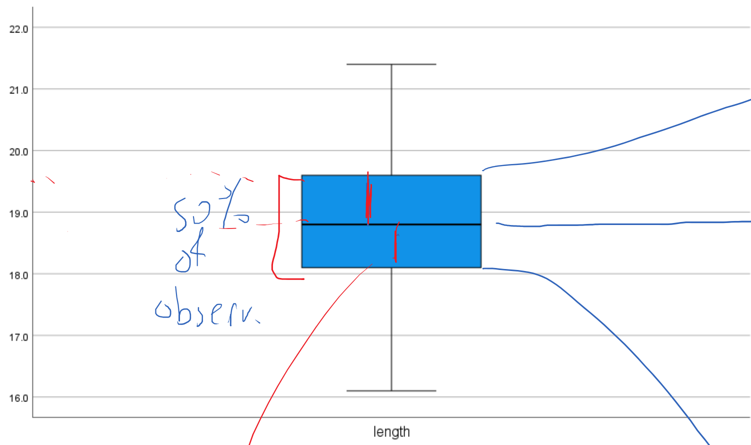
$$CV = \frac{S}{|\bar{X}|} = \frac{1.1205}{18.802}$$

$$= 0.0595 < 0.1$$

Rule

if $CV > 0.1$ (10%) the sample comes from an heterogeneous population

if $CV < 0.1$ (10%) the sampel comes from an homogen. population



50% of observ.

Q_3 — 25% / 75%

median m — 50% ≥ 0 / 50% ≤

Q_1 — 75% / 25%

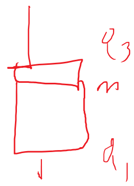
$$M - Q_1 \simeq Q_3 - M$$

Symmetric

$$Q_1 \simeq 18.1$$

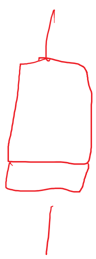
Q_3

Q_3



$-m < m - Q_1$

left skewed



$-m > m - Q_1$

Assume $X \sim N(\mu, \sigma)$

Estimated Distribution Parameters

		length
Normal Distribution	Location	18.802
	Scale	1.1205

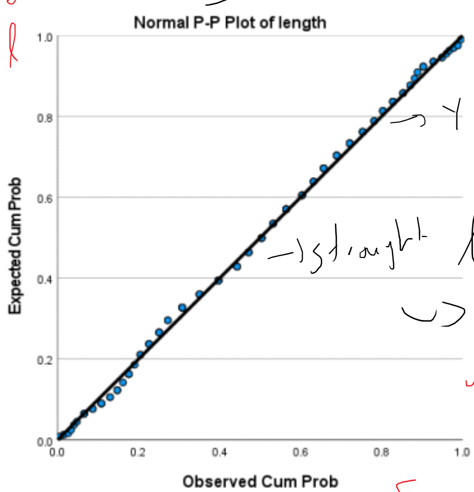
The cases are unweighted.

$\bar{x} = \hat{\mu}$ estimation of μ

$\hat{\sigma}$ estimation of σ

$= s$

theoretical Normal $F(x)$



$\rightarrow y = x$

$\rightarrow$ straight line

$\rightarrow$ Assumption of normal distrib can not be rejected

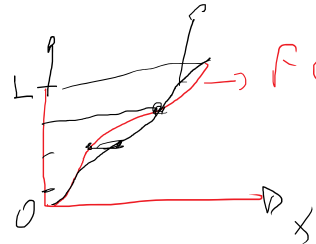
$F(x) = P(X \le x)$

P

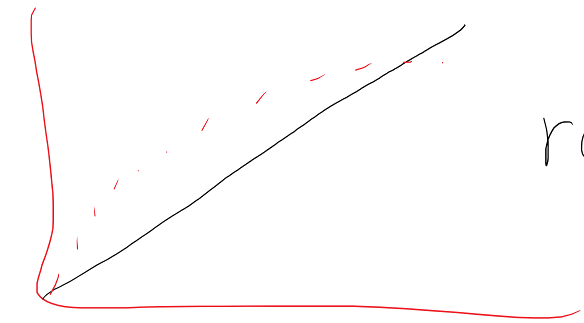
P-P plot

observed $F(x)$

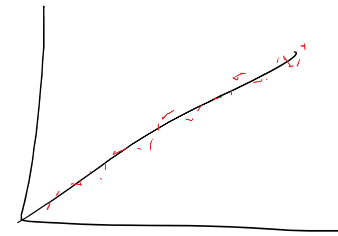
$\rightarrow F(x)$ Norm

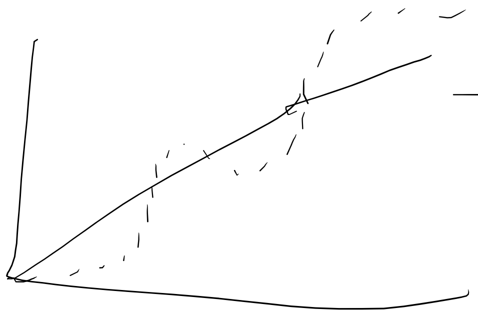


Normal P-P plot

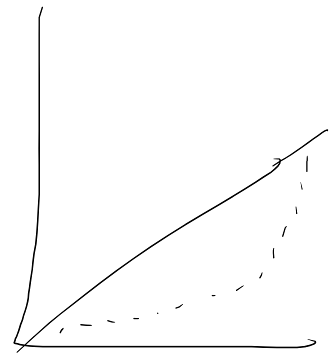


the normal is rejected





→ reject.



reject

is not
rejected

Statistics

		tire	driver
N	Valid	75	75
	Missing	0	0
Mean		1.9733	.5333
Median		2.0000	1.0000
Mode		2.00	1.00

Frequency Table

tire → 3 types

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	25	33.3	33.3	33.3
	2.00	27	36.0	36.0	69.3
	3.00	23	30.7	30.7	100.0
	Total	75	100.0	100.0	

$\hat{p}_1 = 0.333 \to 33.3\%$

$\hat{p}_2 = 0.36 \to 36\%$

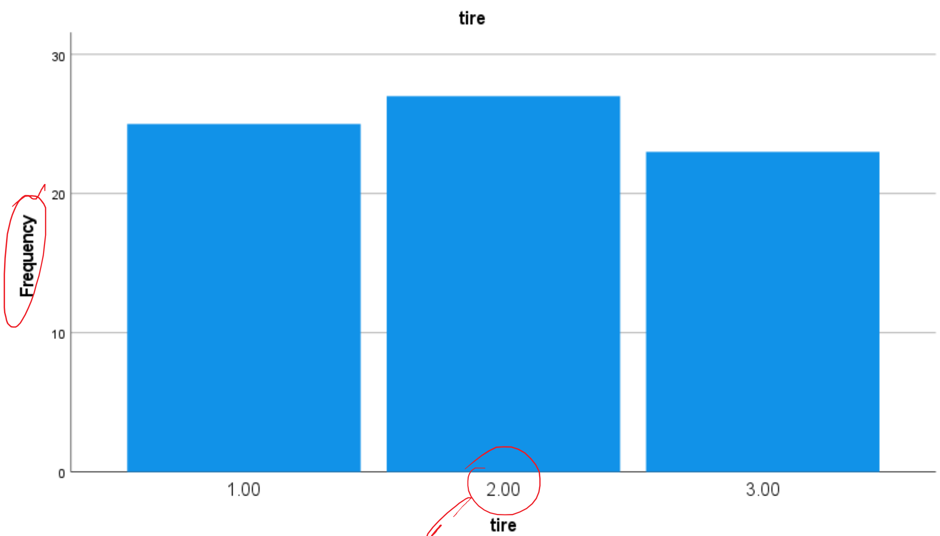
$\hat{p}_3 = 0.307 \to 30.7\%$

driver

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	35	46.7	46.7	46.7
	1.00	40	53.3	53.3	100.0
	Total	75	100.0	100.0	

$\hat{p}_0 = 46.7\%$

$\hat{p}_1 = 53.3\%$



mode

Statistics

tire

		tire
N	Valid	75
	Missing	0
Mean		1.9733
Median		2.0000
Mode		2.00

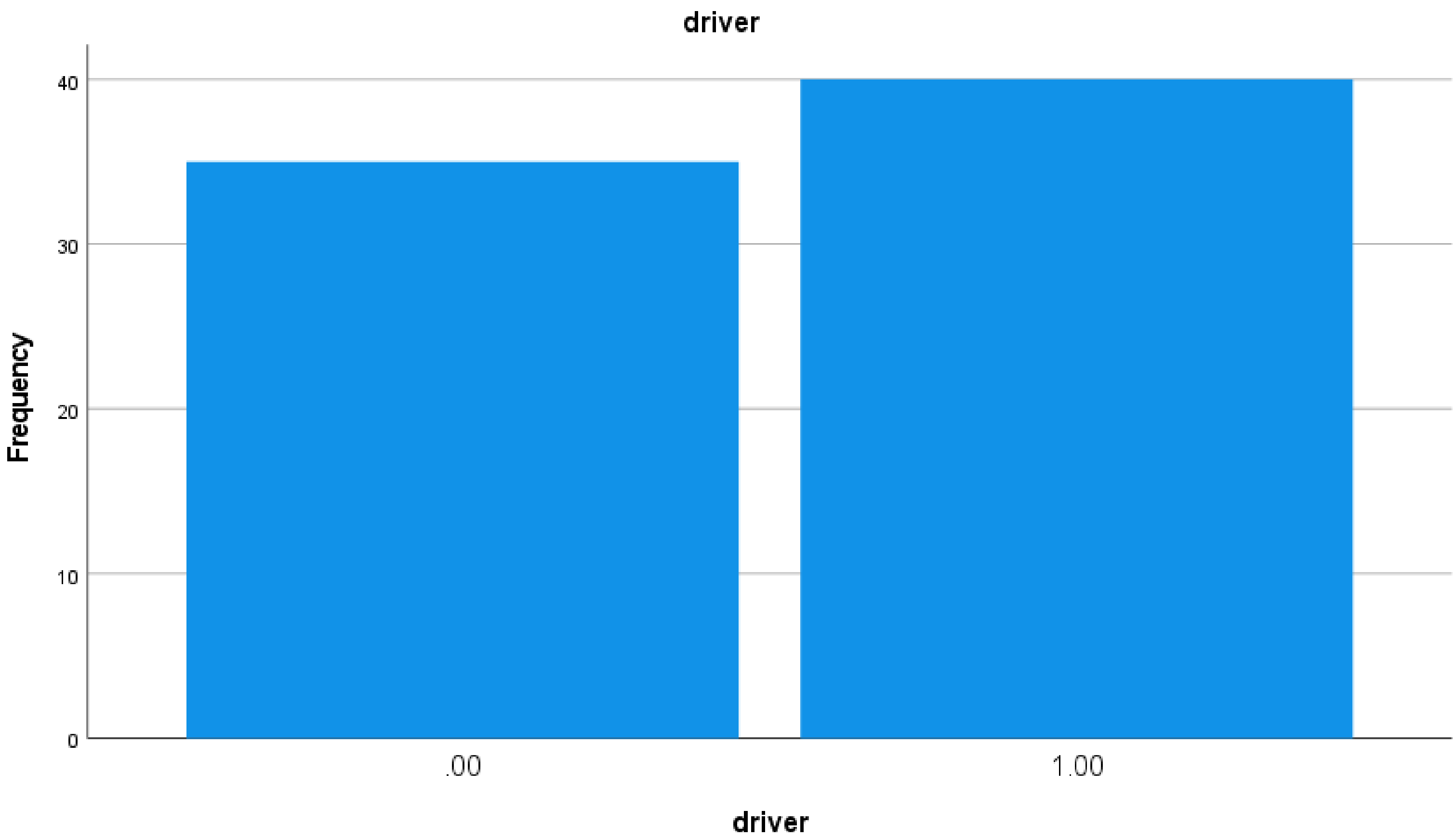
≤ 50%

≥ 50%

$\bar{x} \simeq$ median $\simeq$ mode

symmetric dist.

driver
75
0
.5333
1.0000
1.00



no symmetry

tire * driver Crosstabulation

			driver		
			.00	1.00	Total
tire	1.00	Count	12	13	25
		% within tire	48.0%	52.0%	100.0%
		% within driver	34.3%	32.5%	33.3%
		% of Total	16.0%	17.3%	33.3%
	2.00	Count	11	$\boxed{16}$	27
		% within tire	40.7%	59.3%	100.0%
		% within driver	31.4%	40.0%	36.0%
		% of Total	14.7%	21.3%	36.0%
	3.00	Count	12	11	23
		% within tire	52.2%	47.8%	100.0%
		% within driver	34.3%	27.5%	30.7%
		% of Total	16.0%	14.7%	30.7%
Total		Count	35	$\boxed{40}$	75
		% within tire	46.7%	53.3%	100.0%
		% within driver	100.0%	100.0%	100.0%
		% of Total	46.7%	53.3%	100.0%

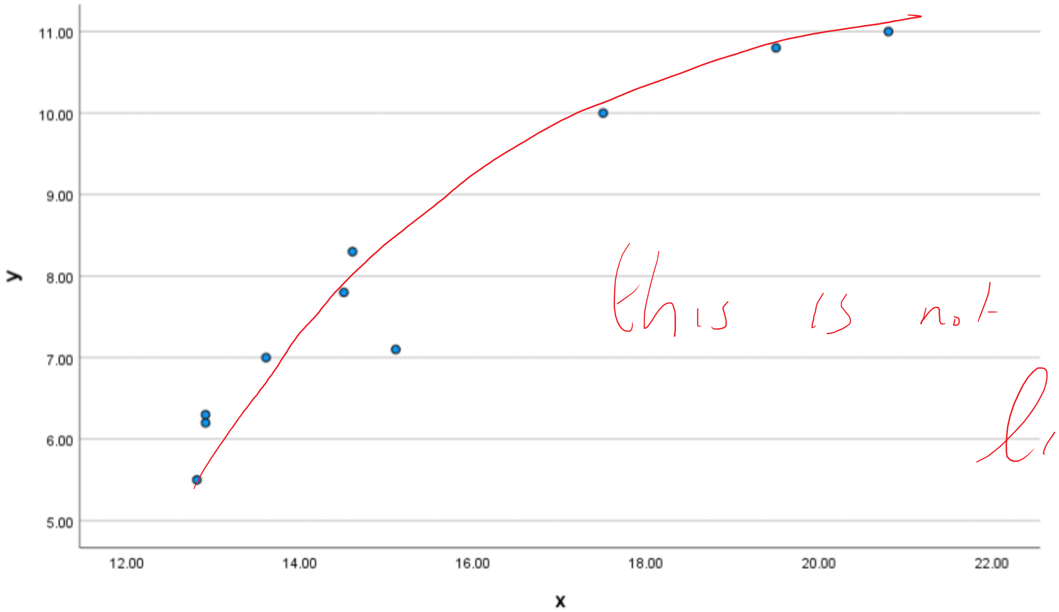
16 times that we used driver 1 with type 2

$$59.3\% = \frac{16}{27}$$

21.3

$$40\% = \frac{16}{40}$$

$$\% = \frac{16}{75}.$$



this is not a linear relat.

x P -> y

MLE

Suppose that we have observed

$x_1, x_2, \ldots, x_n$

from a distribution with probab. function (disc

$f(x; \theta)$

then the probab. of the observed sample is

$$P(X_1 = x_1 \cap X_2 = x_2 \cap \cdots \cap X_n = x_n)$$

(rele)

$$P(X_1=x_1 \cap X_2=x_2 \cap \cdots \cap X_n=x_n) \overset{\text{idep.}}{=}$$

$$P(X_1=x_1)\cdot P(X_2=x_2)\cdots P(X_n=x_n)$$

$$= \underbrace{f(x_1;\theta)\cdot f(x_2;\theta)\cdots f(x_n;\theta)}_{\text{observed values of the sample}} = L(\theta)$$

known, unknown, known, p, known, p

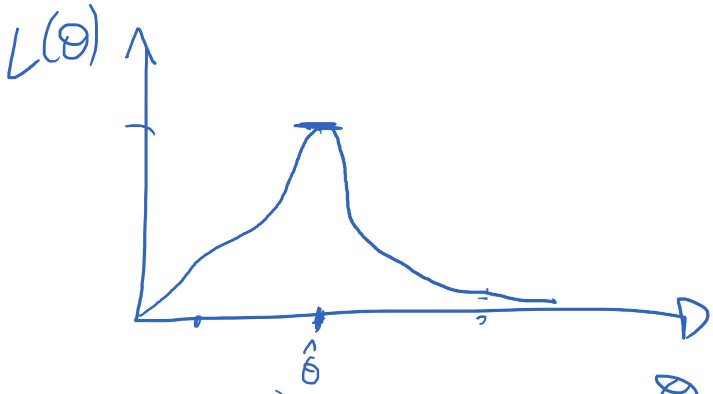
prob. to observe what we have observed

this is a

4un Fire

of

9



point that maxim. $L(\theta)$

$\hat{\theta}$: $\max L(\theta)$

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

$$\ell(\theta) = \log L(\theta)$$

$$= \log \prod f$$

$$= \sum \log f$$

log-likelihood

$$\ell'(\theta) = 0$$

solve

mle $\hat{\theta}$

$(x_i; \theta)$

$(x_j; \theta)$