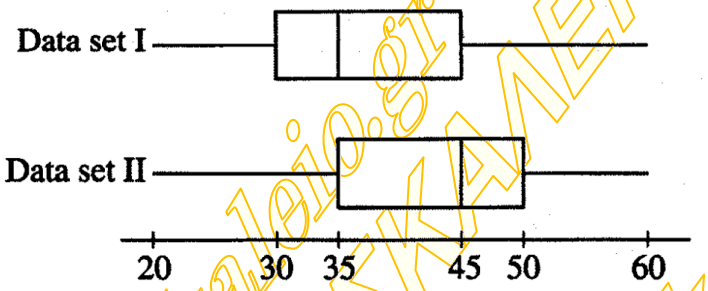
1.5 The boxplots (or Box and Whisker plots) shown below summarize two data sets, I and II. Based on the boxplots, which of the following statements about these two data sets CANNOT be justified?



a. The range of data set I is equal to the range of data set II.
b. The interquartile range of data set I is equal to the interquartile range of data set II.
c. Data set I and data set II have the same number of data points.
d. About 75% of the values in data set II are greater than or equal to about 50% of the values in data set I.

Answer: reduction to the absurd (εις άτοπον επαγωγή)

a. Yes 20 – 60

b. Yes, Data I : 45-30 = 15 and Data II : 50-35 =15

d. Yes Data II Q_1=35 (smaller than 75%) is equal to Data I, Q_2=Median (larger than 50%)

c. Not enough information

1.8 Future parents are wondering how many boys they might get if they have three children. A probability model is developed with _____ **possible outcomes**.

Let s denote B = {boy} and G={ girl} . Having 3 children produces 8 **possible outcomes**

BBB, GBB, BGB, BBG ,GGB ,GBG, BGG, GGG

Now, let s denote X= the discrete variable that counts the number of Boys then

B	0	1	2	3
Probability	1/8 (GGG)	3/8	3/8	1/8

1.9 Let $X_1,..,X_{16}$ be a random sample from a Normal distribution with mean 5 and variance 48. What is the distribution of the sample mean $\overline{X}$?

The distribution is exactly Normal, cause the random sample was selected from a Normal distriburion. If it was not, by using the central limit theory within a large sample selected , it would be approximately Normal.

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \text{ so } \quad mean = 5 \quad \text{var}\, iance = \frac{48}{16} = 3$$

1.20 You have estimated the model Y=1.8*X^(-1.35). If X changes by 10% which of the following statements is true?

a. Y changes by 1.35%
b. Y changes by -13.5%
c. Y changes by +13.5%
d. Y changes by -1.35%

Constant elasticity model so -1,35 is the elasticity which means if X changes by 1% then Y changes by -1,35%. By using linear interpolation (γραμμική παρεμβολή – μέθοδο των τριών) if X changes by 10% then Y changes by -13,5%.

Question 2

The leader of a small political party which participates in the elections wonders if his party will win more than 3% of votes, which is the minimum to win seats in parliament. The last poll based on a sample size of 2,500 voters, gave an estimation of 3.2%. The leader of the party wonders if this percentage is significantly higher than 3%.

2.1 State the null hypothesis (H_0) and the alternative hypothesis (H_1)

2.2 Test the hypothesis at α = 5%.

2.3 Set up a 95% confidence interval of the party's percentage.

2.4 Given that the total number of voters is 9,000,000 and one third does not participate in elections, set up a 95% confidence interval of the number of voter's who are expected to vote for the party.

2.5 What will be your recommendation to the party leader?

Note: $s_p = \sqrt{\frac{p\cdot(1-p)}{n}} = 0.0035$, $Z = (p - \pi_0)/s_p = +0{,}571$, $|Z_{0.10/2}| \approx 1{,}65$, $|Z_{0.05/2}| \approx 2{,}00$

Answer: Point estimation : $\hat{p} = 0,032$ or $3,2\%$

2.1.

$H_0: \pi = 0.03$

$H_1: \pi > 0.03$ (> : The leader of the party wonders if this percentage is significantly higher than 3%.)

2.2 Hypothesis testing :

$$z = \frac{0.032 - 0.03}{\sqrt{\frac{0.032(1-0.032)}{2.500}}} = \frac{0.002}{3.52\cdot 10^{-3}} = \frac{0.002}{0.00352} = 0.568$$

$so, |z| = 0,568$

• Ho is rejected if $|z| > z_a = z_{0.05} = 1,65$

While $|z| = 0,568 < z_{0.05} = 1,65$ Ho is not rejected, so the percentage **is not** significantly higher than 3%.

2.3 : 95% C.I. :

$$\left(p - z_{\alpha/2}\cdot S\hat{E}(\hat{p}) \;,\; \hat{p} + z_{\alpha/2}\cdot S\hat{E}(\hat{p})\right) \Rightarrow (0.032 - 2\cdot 0.00352 \;,\; 0.032 + 2\cdot 0.00352)$$

$$\Rightarrow (0.025 \;,\; 0.039)$$

2.4. Number of voters that participate in the elections :

$$N = 9.000.000 - \frac{1}{3}\cdot 9.000.000 = 6.000.000$$

so 95% C.I : $(6.000.000\cdot 0.025 \;,6.000.000\cdot 0.039) = (150.000,23\;\; 4.000)$

2.5 . The estimated percentage leads to conclusion that is not significantly higher than 3% (at α = 5%). However, it is not certain (like **everything** in statistics..) that the party will enter parliament. But it is likely to achieve it.