Problemas: 10

Lista de 10 problemas aprox y haremos en clase unos 3 o los que de tiempo

**CASE STUDY**

Imagen que contiene Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

**DATA BASICS**

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

**Data collection principles**

Texto

El contenido generado por IA puede ser incorrecto.

Texto

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

Imagen que contiene Texto

El contenido generado por IA puede ser incorrecto.

Texto

El contenido generado por IA puede ser incorrecto.Tabla

El contenido generado por IA puede ser incorrecto. value of a quantity and its true value. In statistics, an error is not a "mistake." Variability is an inherent part of the results of measurements and of the measurement process.

Observed errors can be divided into two components: random error and systematic error.

Systematic errors are errors that are not determined by chance but are introduced by an inaccuracy (involving either the observation or measurement process) inherent to the system. Random errors are related to the sampling. Each subsequent measurement has a random error, leading to imprecision in the estimation. A measurement with a low random error is said to be precise.

In systematic error each subsequent measurement has the same recurring error do to a bias.

Four analysts, A, B, C and D, each prepared five replicate samples to measure the pH of a specific sample of soil. Results are following

Mean SD

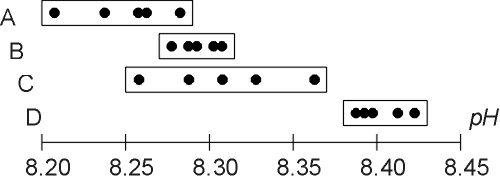
A 8.208 8.239 8.258 8.264 8.283 8.250 0.028

B 8.278 8.288 8.293 8.304 8.308 8.294 0.012

C 8.259 8.289 8.308 8.329 8.363 8.310 0.039

D 8.389 8.393 8.399 8.413 8.423 8.403 0.014

The four sets of results are shown diagrammatically below:



1. Which set has less random error? which has the highest random error?
2. Which set is more precise? which set is less precise?
3. Is any of the sets likely to have a systematic error (to be biased)?
4. If the true value were known to be 8.31, which set would be more accurate?
5. Look for a definition of precision and accuracy in statistics.

Exercise 5. In a certain population of the freshwater sculpin, Cottus rotheus, the distribution of the number of tail vertebrae, Y, is as shown in next Table

Tabla

El contenido generado por IA puede ser incorrecto.

a. Calculate the mean of Y and the variance of Y

remember the mean of a discrete random variable Y is defined as and the variance of a discrete random variable Y is defined as

Exercise 4.

Random error is function of the standard deviation. B has less random error and C has the highest random error

1. Which set is more precise? which set is less precise?

Lower standard deviation indicated higher precision. B is more precise, and C is less precise

1. Is any of the sets likely to have a systematic error (to be biased)?

D gives a divergent result; D is likely to be biased

1. If the true value were known to be 8.31, which set would be more accurate?

C estimates is equal to the true values. Then C would be the more accurate. In this example the less precise is the more accurate.

Exercise 5.

The mean of Y is 21.49 and the variance of Y is 0.4299

Exercise 3. Following are data on the number of virus resistant bacteria on an experimental study

14 15 13 21 15 14 26 16 20 13

a. Determine the median and the quartiles

b. Determine the interquartile range

c. How large would an observation in this data set have to be in order to be an outlier?

a. Sort the data

13 13 14 14 15 15 16 20 21 26

First we determine the median. n = 10 is an even number, the median is the number located in the position 0.5x(n+1) = 5.5

The value at position 5 is 15, and the value at position 6 is 15. The average of both values is 15, then the median is 15.

13 13 14 14 15 (15) 15 16 20 21 26

The first quartile is the median of the first 50% of the data, given is an odd value, the first quartile is 14

In the same way, the third quartile is 20

b. interquartile range is Q3 – Q1 = 20-14 = 6

c. to determine the outliers we need to determine the lower fence and the upper fence

Lower fence = Q1 – 1.5xIQR = 14 – 1.5 x 6 = 14 – 9 = 5

Upper fence = Q3 + 1.5 x IQR = 20 + 1.5 x 6 = 20 + 9 = 29

To be considered an outlier, a value must be lower than 6 or higher than 29.

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Texto

El contenido generado por IA puede ser incorrecto.

Texto

El contenido generado por IA puede ser incorrecto.