

ANÁLISIS DE DATOS – ESTADÍSTICA DESCRIPTIVA

Estadística descriptiva

- Procedimientos y medidas que describen datos cuantitativos.
- Distribución de frecuencias – histogramas
- Descripción de distribución de resultados:
 - ▣ Ubicación, dispersión y forma
 - ▣ Ubicación: medidas de tendencia central (media, mediana, moda)
 - ▣ Dispersión: medidas de variabilidad (desviación estándar, varianza, rango)
 - ▣ Forma de la distribución: Normal, Poisson, alguna otra.

Medidas de tendencia central

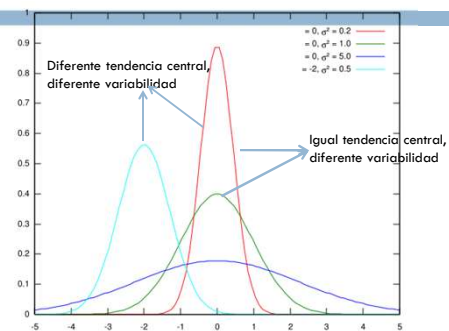
- Media, mediana, moda – Cómo está en general la muestra en relación a la variable medida.

```
> summary(misdatos)
```

SEXO	ESTATURA	PESO	EDAD	PROMEDIO
Min. :0.0000	Min. :1.150	Min. : 80.0	Min. :17.00	Min. :6.000
1st Qu.:0.0000	1st Qu.:1.600	1st Qu.:120.0	1st Qu.:20.00	1st Qu.:7.140
Median :1.0000	Median :1.680	Median :138.0	Median :21.00	Median :7.490
Mean :0.5175	Mean :1.666	Mean :137.7	Mean :21.58	Mean :7.524
3rd Qu.:1.0000	3rd Qu.:1.720	3rd Qu.:155.0	3rd Qu.:23.00	3rd Qu.:7.942
Max. :1.0000	Max. :1.970	Max. :235.0	Max. :29.00	Max. :9.450

- Máximos y mínimos y cuartiles idea de distribución

Distribuciones

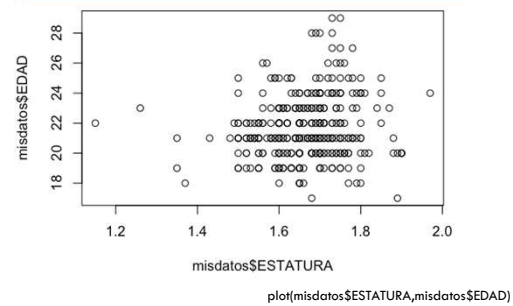


Dispersión

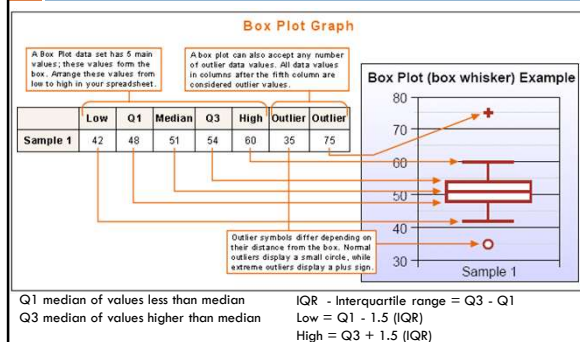
- Desviación Estándar
- Varianza
- Gráficos de dispersión

```
> sd(misdatos$ESTATURA)      > var(misdatos$PESO)
[1] 0.09513567                [1] 595.7699
> sd(misdatos$PESO)
[1] 24.4084
```

Gráficos de dispersión



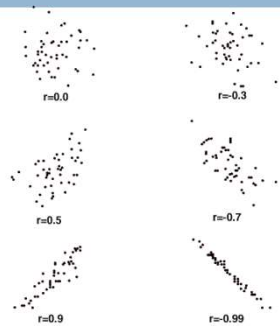
Qué significan las líneas y el gráfico



Correlaciones

- El grado de asociación entre dos variables se llama correlación => coeficiente de correlación.
- Toma valores entre -1.00 y 1.00
- Los diagramas de dispersión, nos muestran de forma gráfica esa correlación.
- Mientras más alto es el valor del coef. De correlación, más cercanos están los puntos de la línea que los describe.
- Usado para describir relaciones entre 2 variables y predecir el resultado de una variable en función de otra.

Algunos ejemplos de diagramas de dispersión



Correlaciones y fortaleza

Coefficiente de correlación	Fortaleza de la relación
0.7 – 1.0	Fuerte
.30 - .69	Moderada
+0.0 - .29	Ninguna o débil

Taller

- Leer, discutir y resumir paper distribuido en clase.

Tarea

- Instalar R Studio en sus computadoras. Traer al menos una laptop por grupo para la proxima clase.

