Quantity of information:
1 octet (byte) (symbol o / B) = 8 bits

Symbol	Binary	Byte
kbyte (kilobit) – kB	2^{10}	1024
Mbyte (megabit) – MB	2^{20}	1048576
Gbyte (gigabit) – GB	2^{30}	1073741824
Tbyte (terabit) – TB	2^{40}	1099511627776

DESCRIPTIVE STATISTICS

Median:

n = odd

$$Me = x_{\frac{n+1}{2}}$$

n = even

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

Mean:

Population

$$\mu = \frac{\sum_{i=1}^{n} X_i}{N}$$

Sample

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Skewness

$$M_3 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^3}{n}$$

Kurtosis

$$\alpha_4 = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^4}{S^4} - 3$$

Amplitude (Range): $R = X_{max} - X_{min}$

$$Centralvalue = \frac{X_{min} + X_{max}}{2}$$

Means of deviations:

From the mean

$$R_{\bar{X}} = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n}$$

From the median

$$R_{Me} = \frac{\sum_{i=1}^{n} |X_i - Me|}{n}$$

Variance:

Population

$$\sigma^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{N}$$

Sample

$$s^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}$$

Standard deviation:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}}$$

Standard error (ES):

$$ES = \frac{s}{\sqrt{n}}$$

Coefficient of variance (CV):

$$CV = \frac{s}{\bar{X}}$$

PROBABILITIES

Pr(A) = (number of favourable cases)/(number of possible cases)

Sensibility: Se = Pr(B|A)

where Pr(B) = probability of a positive test; Pr(nonB) = probability of a no positive test; Pr(A) = probability of existence of a certain disease; Pr(nonA) = probability of non existence of a certain disease.

Specificity: Sp = Pr(nonB|nonA)

Positive predictive value: PPV = Pr(A|B)

Negative predictive value: NPV = Pr(nonA|nonB)

False positive rate: FPR = Pr(B|nonA)

False negative rate: FNR = Pr(nonA|B)

Addition rules: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$

Addition rule for mutually exclusive events:

$Pr(A \cup B) = Pr(A) + Pr(B)$

Multiplication rules: $Pr(A \cap B) = Pr(A) \cdot Pr(B|A)$

Multiplication rule for independent events:

$Pr(A \cap B) = Pr(A) \cdot Pr(B)$

BAYES Formula:

$$Pr(A|B) = \frac{Pr(B|A) \cdot Pr(A)}{Pr(B|A) \cdot Pr(A) + Pr(B|nonA) \cdot Pr(nonA)}$$

RANDOM VARIABLE

Mean (expected value):

$$M(X) = \sum_{i=1}^{n} X_i \cdot Pr(X_i)$$

Variance:

$$V(X) = \sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)$$

Standard deviation:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)}$$

BINOMIAL RANDOM VARIABLES

Binomial DISTRIBUTION:

$$Pr(X = k) = C_n^k p^k q^{n-k} \text{ where } C_n^k = \frac{n!}{k!(n-k)!}$$

Mean: $M(X) = n \cdot p$

Variance: $V(X) = n \cdot p \cdot q$

Standard deviation: $\sigma(X) = \sqrt{V(X)} = \sqrt{n \cdot p \cdot q}$

POISSON RANDOM VARIABLE

$$X : \begin{pmatrix} k \\ e^{-\theta} \cdot \frac{\theta^k}{k!} \end{pmatrix} \qquad Pr(X = k) = \frac{e^{-\theta} \cdot \theta^k}{k!}$$

e = 2.718281828; θ = n·p (n = sample size, p = probability of the apparition of an event).

CONFIDENCE LEVELS

Mean:

$$\left[\bar{X} - Z_\alpha \frac{s}{\sqrt{n}}, \bar{X} + Z_\alpha \frac{s}{\sqrt{n}}\right]$$

Frequency:

$$\left[f - Z_\alpha \sqrt{\frac{f \cdot (1-f)}{n}}; f + Z_\alpha \sqrt{\frac{f \cdot (1-f)}{n}}\right]$$

PEARSON CORRELATION COEFICIENT

$$r = \frac{\sum\left(X - \bar{X}\right)\left(Y - \bar{Y}\right)}{\sqrt{\sum\left(X - \bar{X}\right)^2 \sum\left(Y - \bar{Y}\right)^2}}$$

CHI-SQUARE TEST (χ2): $\chi^2 = \sum_{i=1}^{r\cdot c} \frac{(f_i^0 - f_i^t)^2}{f_i^t}$

where χ^2 = χ^2 test statistic; f_i^o = observed frequency; f_i^t = theoretical frequency.

Critical region for α = 0.05 is [3.84, ∞).

- If $\chi^2 \in [3.84; \infty)$ H_0 is not accepted with a type I risk of error (α).
- If $\chi^2 \notin [3.84; \infty)$ H_0 is accepted with a type II risk of error (β).

Z TEST FOR PROPORTIONS (comparison of an observed frequency with a theoretical frequency)

$$z = \frac{f - p}{\sqrt{\frac{p(1-p)}{n}}}$$

where p = theoretical frequency (population frequency); f = observed frequency, n = sample size.

Z TEST FOR PROPORTIONS (testing the equality of two proportions)

$$z = \frac{(p_1 - p_2)}{\sqrt{p\cdot(1-p)\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}, \quad p = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$$

where p_1 = frequency for 1st sample; n_1 = sample size of 1st sample; p_2 = frequency of 2nd sample; n_2 = sample size for 2nd sample.

Z TEST FOR COMPARING OF A SAMPLE MEAN WITH A POPULATION MEAN (equal variances)

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

where μ = population mean; $\bar{X}$ = sample mean; σ = population standard deviation; n = sample size

Critical region for α = 0.05 (two-tailed test): (-∞ , -1.96] ∪ [1.96 , ∞)

RISKS AND ODDS ON 2×2 CONTINGENCY TABLE

	Disease+	Disease -	Total
Test +	AP	FP	= AP+FP
Test -	FN	AN	= FN+AN
Total	= AP+FN	=FP+AN	= AP+FP+FN+AN = n

Name	Formula
False positive rate	=FP/(FP+AN)
False negative rate	=FN/(FN+AP)
Sensibility	=AP/(AP+FN)
Specificity	=AN/(AN+FP)
Accuracy	=(AP+AN)/n
Predictive positive value	=AP/(AP+FP)
Predictive negative value	=AN/(AN+FN)
Relative risk	=AP(FP+AN)/FN(AP+FP)
Odds ratio	=(AP·AN)/(FN·FP)
Attributable risk	=AP/(AP+FP)-FN/(FN+AN)

STUDENT (T) TEST FOR COMPARING A MEAN WITH A KNOWN MEAN (unknown variances)

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

where μ = population mean; $\bar{X}$ = sample mean; s = sample standard deviation; n = sample size.

- **Degree of freedom (df)*:*** df = n-1
- **Critical region** for α = 0.05 (two-tailed test): $(-\infty, -t_{n-1, 0.025}] \cup [t_{n-1, 0.025}, \infty)$

STUDENT (T) TEST FOR COMPARING TWO MEANS (equal variances)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s\cdot\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}, \text{where } s = \sqrt{\frac{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1+n_2-2}}$$

$\bar{X}_1$ = mean of 1st sample; n_1 = sample size of 1st sample; s_1^2 = variance of 1st sample; $\bar{X}_2$ = mean of 2nd sample; n_2 = sample size of 2nd sample; s_2^2 = variance of 2nd sample

STUDENT (T) TEST FOR COMPARING TWO PAIRED MEANS

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}, \text{where } \bar{d} = \frac{d_1 + d_2 + ...d_n}{n}$$

s = standard deviation of differences; n = sample size

Z TEST FOR COMPARIND THE MEANS OF TWO POPULATIONS (KNOWN AND UNEQUAL VARIANCES)

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $\bar{X}_1$ mean of 1st sample; n_1 = sample size of first sample; s_1^2 = variance of first sample; $\bar{X}_2$ = mean of 2nd sample; n_2 = sample size of 2nd sample; s_2^2 = variance of 2nd sample.

SAMPLE SIZE ESTIMATION: MEAN TESTING (normal distribution) / SAMPLE SIZE ESTIMATION: NO OBJECTIVE PRIOR DATA

$$n = \frac{(z_{1-\alpha} - z_{1-\beta})^2 \sigma^2}{(\bar{X} - \mu)^2}$$

Critical value two-tailed test: $z_{1\text{-}5\%}$ = 1.960, $z_{1\text{-}\beta(\beta=20\%)}$ = 0.842

SAMPLE SIZE ESTIMATION: TEST OF TWO MEANS (normal distribution)

$$n_1 = n_2 = \frac{(z_{1-\alpha/2} - z_{1-\beta})^2 (\sigma_1^2 + \sigma_2^2)^2}{d^2}$$

SAMPLE SIZE ESTIMATION: MEAN (non normal distribution)

$n = \frac{\sigma^2}{\alpha \cdot k^2}$, where k = the difference you want to detect between the sample mean and population mean (a clinical choice)

SAMPLE SIZE ESTIMATION: ON RATES

Central proportion (is not near 0 or 1):

$$n = \left[\frac{z_{1-\alpha/2}\sqrt{\pi\cdot(1-\pi)} + z_{1-\beta}\sqrt{p\cdot(1-p)}}{p - \pi}\right]^2$$

. where π = theoretical proportion; p = desired proportion

Extreme proportion (is near 0 or 1):

$$n = \left(\frac{z_{1-\alpha/2}\sqrt{\pi} + z_{1-\beta}\sqrt{p}}{p - \pi}\right)^2$$