misc

•
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

• Bayes:
$$P(B|A) = \frac{P(A|B)P(B)}{P(B)}$$

•
$$(S^*)^2 = \frac{n}{n-1}S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2;$$

•
$$S^* = \sqrt{(S^*)^2}, S = \sqrt{S^2}$$

•
$$\operatorname{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$
;

•
$$cov(X, Y) = \mathbb{E}\left((X - \bar{X})(Y - \bar{Y})\right) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y);$$

ordinary least squares

$$b = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2} = \frac{\sum xy}{x^2},$$

$$x = X - \bar{X}, y = Y - \bar{Y}$$

 $a \text{ from } Y = a + bX$
(the line goes to through (\bar{X}, \bar{Y}))

standard error:

the correct regresion line is $Y_i = \alpha_i + \beta_i X_i + \varepsilon_i$

Standard error:

$$SE = \frac{\sigma}{\sqrt{\sum x^2}} = \frac{\sigma}{\sqrt{n}S_x}$$

Smol SE is good.

standard squared error

(we lost 2 degs of freedom)

$$SSE = \frac{1}{n-2} \sum (Y - \hat{Y})^2$$

P values and stuff

$$t = \frac{b}{SE}$$

 H_0 : b = 0 (ie there's no relation) if $t > T_{\alpha}^{(n-2)}$ then we reject H_0 else we don't (with α liek .005)

F-statistic

(dont worry about it) similar to T (from above, but with)

$$F = \frac{\text{RegSS}}{\text{SSE}/(n-2)}$$

and reject if $> F_{\alpha}^{(1,n-2)}$

 \mathbb{R}^2

$$R^{2} = \frac{\text{RegSS}}{\text{TSS}} = 1 - \frac{SSE}{TSS} = 1 - \frac{\sum (Y - \bar{Y})^{2}}{\sum (Y - \hat{Y})^{2}}$$

 \hat{Y} is the esit mated value of Y (ie $\hat{Y}=a+bX)$ RegSS $=\sum (\hat{Y}-\bar{Y})^2$

Confidence intervals - tl;dr (this is a lot more information than we need) for the average

type	useful	X type	n size	σ known	the interval	
bilateral	X	$\mathcal{N}(\mu, \sigma^2)$	whatever	yes	$\mu \in \left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right)$	
bilateral	X	whatever	big	yes	$\mu \in \left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right)$	
no sup		$\sim \mathcal{N}(\mu, \sigma^2)$	big	yes	$\mu \in \left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}, \infty\right)$	
no inf		$\sim \mathcal{N}(\mu, \sigma^2)$	big	yes	$\mu \in \left(-\infty, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right)$	
bilateral	X	whatever	big	no	$\mu \in \left(\bar{X} - \frac{S^*}{\sqrt{n}} z_{1-\alpha/2}, \bar{X} + \frac{S^*}{\sqrt{n}} z_{1-\alpha/2}\right)$	
unilateral		whatever	big	no	like rows 2 and 3 but with S^*	
bilateral	X	$\sim \mathcal{N}(\mu, \sigma^2)$	smol	no	$\mu \in \left(\bar{X} - \frac{S^*}{\sqrt{n}} t_{1-\alpha/2, n-1}, \bar{X} + \frac{S^*}{\sqrt{n}} t_{1-\alpha/2, n-1}\right)$	
unilateral		$\sim \mathcal{N}(\mu, \sigma^2)$	smol	no	like rows 2 and 3 but with S^* and $t_{1-\alpha,n-1}$	

for the variance (dispersion)

just ignore this

type	where to find	X type	n size	μ known	the interval	
bilateral	C5 - pg 6	$\sim \mathcal{N}(\mu, \sigma^2)$	smol	yes	$\sigma^2 \in \left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\alpha/2,n}^2}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\alpha/2,n}^2}\right)$	
bilateral	C5 - pg 6	$\sim \mathcal{N}(\mu, \sigma^2)$	smol	no	$\sigma^2 \in \left(\frac{(n-1)(S^*)^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)(S^*)^2}{\chi^2_{1-\alpha/2,n-1}},\right)$	

for n large, look at C5 pg 7 - obs 1.3 tl;dr we make it a normal distribution

$Testarea\ ipotezelor\ statistice\ -\ tl; dr$

don't worry about it

For the average

 $(H_0): \mu = \mu_0$

name	where to find	X type	n size	σ known	$thing_0$	bilateral tl;dr
Z test	C6 - pg 6-9	$\sim \mathcal{N}(\mu, \sigma^2)$	big	yes	$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	$z_0 \in \left(-z_{1-\alpha/2}, z_{1-\alpha/2}\right)$
T test	C6 - pg 10-13	$\sim \mathcal{N}(\mu, \sigma^2)$	smol	no	$z_0 = \frac{\bar{x} - \mu_0}{s^* / \sqrt{n}}$	$t_0 \in \left(-t_{1-\alpha/2,n-1}, t_{1-\alpha/2,n-1}\right)$

For the variance (dispersion)

$$(H_0): \sigma = \sigma_0$$

$$\chi^2\text{-test}$$

$$\chi_0^2 = \frac{(n-1)(s^*)^2}{\sigma_0^2}$$

 H_0 is acepted (or pedantically "not rejected") if: $\chi_0^2\in(\chi_{1-\alpha/2,n-1}^2,\chi_{1-\alpha/2,n-1}^2)$

Sun Tzu said: "All exams are based on deception". (the quote in statistics was better)