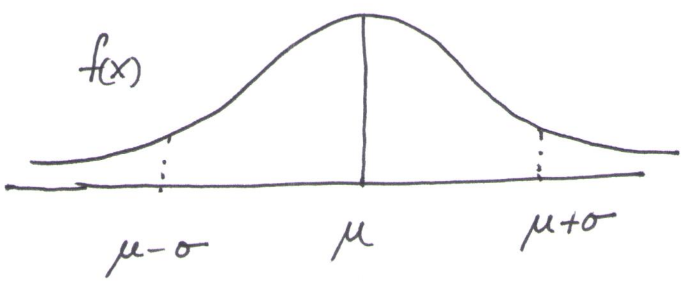
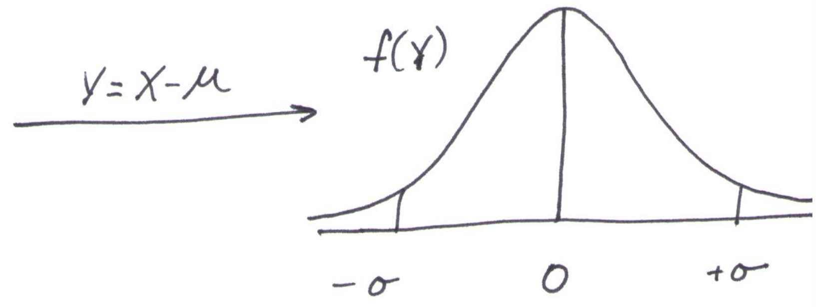
Understanding the Gaussian $1-\alpha$ Confidence Interval

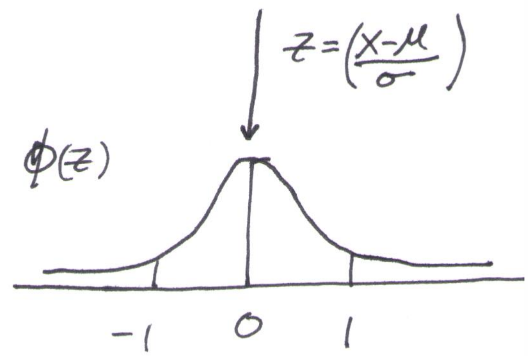
If $X \sim N(\mu, \sigma^2)$ then $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$. Z is the standard Gaussian random variable and we write

$$f(z) = \phi(z) = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu)^2\right\}.$$





Any Gaussian (random variable) density can be rescaled into a standard Gaussian (random variable) density.



Confidence Interval for μ

$x_1, \dots, x_n \quad x_i \sim f(x_i | \mu, \sigma^2) \equiv N(\mu, \sigma^2)$ Assume σ^2 is known. Given $\bar{X}$ the ML and MM estimate of μ find a $1-\alpha$ confidence interval for μ

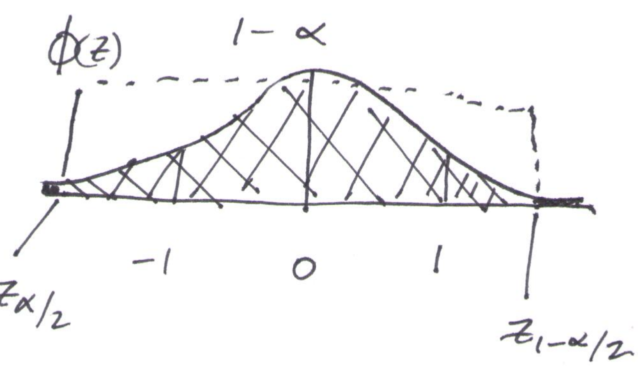
$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{n^{\frac{1}{2}}(\bar{X}-\mu)}{\sigma} \sim N(0, 1)$$

Pick $\alpha \in (0, 1)$ For example, $\alpha = 0.05$ or $\alpha = 0.01$

$$1-\alpha = Pr(z_{\alpha/2} \le Z \le z_{1-\alpha/2})$$

$$= Pr\left(z_{\alpha/2} \le \frac{n^{\frac{1}{2}}(\bar{X}-\mu)}{\sigma} \le z_{1-\alpha/2}\right)$$

$$= Pr\left(\frac{z_{\alpha/2}\sigma}{n^{\frac{1}{2}}} \le (\bar{X}-\mu) \le \frac{z_{1-\alpha/2}\sigma}{n^{\frac{1}{2}}}\right)$$



$$= Pr\left(-\bar{X} + \frac{z_{\alpha/2}\sigma}{n^{\frac{1}{2}}} \le -\mu \le -\bar{X} + \frac{z_{1-\alpha/2}\sigma}{n}\right)$$

$$= Pr\left(\bar{X} - \frac{z_{\alpha/2}\sigma}{n^{\frac{1}{2}}} \geq \mu \geq \bar{X} - \frac{z_{1-\alpha/2}\sigma}{n}\right)$$

$$= Pr\left(\bar{X} - \frac{z_{1-\alpha/2}\sigma}{n} \leq \mu \leq \bar{X} + \frac{z_{1-\alpha/2}\sigma}{n}\right)$$

$z_{\alpha/2} = z_{1-\alpha/2}$

or

$z_{1-\alpha/2} = -z_{\alpha/2}$

For example if $\alpha = 0.05$

$z_{1-\alpha/2} = z_{0.975} = 1.96$

$z_{\alpha/2} = z_{0.025} = -1.96$

The probability $1-\alpha$ is interpreted in the long-run frequency sense.

Hence, if we have $\hat{\theta}_{ML} \approx N(\theta, I(\theta)^{-1})$ then an approximate 95% CI for θ is

$$\hat{\theta}_{ML} \pm 1.96\, I(\theta_{ML})^{-\frac{1}{2}}$$

where we estimate $I(\theta)$ by $I(\hat{\theta}_{ML})$ (the observed Fisher info).

What if σ^2 is not known?

Estimate σ^2 as s^2 (by convention)

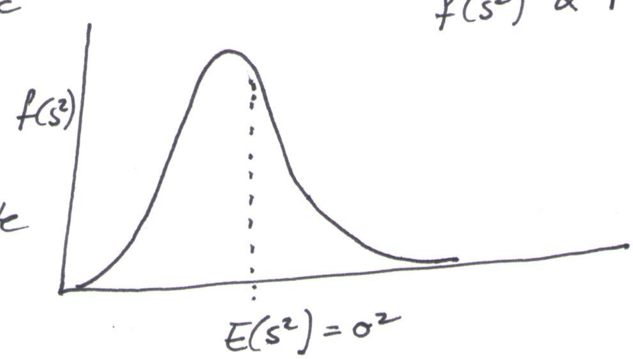
$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$s^2 = g(x_1, \ldots, x_n)$ is a random variable and hence s^2 has a density function $f(s^2)$, a mean $E(s^2)$, a variance $Var(s^2)$ ... etc

$f(s^2) \propto \Gamma(\alpha, \beta)$

$E(s^2) = \sigma^2$ (Unbiasedness)

$$\Rightarrow E(\hat{\sigma}^2) = E\left(\frac{n-1}{n} s^2\right) = \left(\frac{n-1}{n}\right)\sigma^2$$



Because of the importance (desire) of having unbiasedness for the estimate of σ^2, we estimate σ^2 by s^2 and not $\hat{\sigma}^2$, the ML estimate

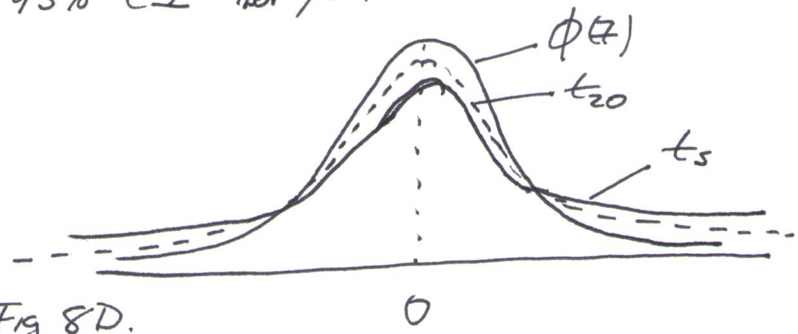
$$\frac{n^{\frac{1}{2}}(\bar{X} - \mu)}{s} \sim t_{n-1}$$

t-distribution with $n-1$ degrees of freedom

$$t_{n-1} \longrightarrow \phi(z)$$

Therefore, the approximate 95% CI for μ is

$$\bar{X} \pm \frac{t_{n-1}(0.975)\, s}{n^{\frac{1}{2}}}$$



n	0.975
10	2.23
20	2.1
25	2.06
30	2.04
60	2.00
120	1.98
∞	1.96

See Lecture 8, Fig 8D.

Confidence Intervals for Regression Coefficients

Give $\hat{\alpha}$ and $\hat{\beta}$, we estimate σ^2 for the regression noise as

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}$$

and $E(s^2) = \sigma^2$

$\Rightarrow$ $Var(\hat{\beta})$ is estimated as $\hat{\sigma}^2_{\beta} = \dfrac{s^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

$Var(\hat{\alpha})$ is estimated as $\hat{\sigma}^2_{\alpha} = \left[\dfrac{1}{n} + \dfrac{\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right] s^2$

$$\frac{\hat{\beta}}{\hat{\sigma}_{\beta}} \sim t_{n-2} \qquad \frac{\hat{\alpha}}{\hat{\sigma}_{\alpha}} \sim t_{n-2}$$

$\Rightarrow$ 95% CI are

$$\hat{\beta} \pm t_{n-2}(0.975)\, \hat{\sigma}_{\beta} \qquad \hat{\alpha} \pm t_{n-2}(0.975)\, \hat{\sigma}_{\alpha}$$

9.073 Simple Regression Addendum 03/06/15 4/5

Analysis of Variance Table

TSS = ESS + RSS (Pythagorean Relationship)

$$\sum_{i=1}^{n}(y_i-\bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2 + \sum_{i=1}^{n}(y_i-\hat{y}_i)^2$$

$$R^2 = \frac{ESS}{TSS}$$ Fraction of the Variance Explained by the Regression Line

$0 \leq R^2 < 1$

ANOVA Table

SS	df	MS	F statistic	p-value
ESS	1	ESS/1	ESS/1/RSS/n-2	
RSS	n-2	RSS/n-2 (s^2)		
TSS	n-1	TSS/n-1		

The p-value is for the null hypothesis $H_0: \beta_0 = 0$ vs $H_A\ \beta \neq 0$ (no slope!)

Parameter estimate	s.e.	t-statistic
$\hat{\alpha}$	$\hat{\sigma}_\alpha$	$\frac{\hat{\alpha}}{\hat{\sigma}_\alpha}$
$\hat{\beta}$	$\hat{\sigma}_\beta$	$\frac{\hat{\beta}}{\hat{\sigma}_\beta}$

For β

$$\boxed{t_{n-2}^2 = F_{1,n-2}}$$

$$R^2 = \frac{F\left(\frac{p-1}{n-p}\right)}{1+F\left(\frac{p-1}{n-p}\right)}$$ $p=2$ here

The R^2 is easier to interpret whereas the F-statistic is easier to work with mathematically because ESS and RSS are independence!

Simple Regression Addendum 03/06/17

ANOVA Table for the Hursh Data

SS	df	MS	$F_{1,65}$	p-value
65,041	1	65,041	1,893	$\ll 0.01$
2,233	65	34.35		
67,274	66			

	SE	t-statistic
$\hat{\alpha} = -3.77$	1.44	-2.61
$\hat{\beta} = 6.07$	0.14	43.5

$R^2 = 0.967$ and $n = 67$ $p = 2$

$$\frac{F \cdot \frac{p-1}{n-p}}{1 + F \cdot \frac{p-1}{n-p}} = \frac{1{,}893 \cdot \frac{1}{65}}{1 + 1{,}893 \frac{1}{65}} = \frac{29.12}{30.12} = 0.967 = R^2$$

$$t_{65}^2 = 43.5^2 = 1{,}893 = F_{1,65}$$

Plots (Diagnostic)

1. Y vs X data

2. $Y - \hat{Y}$ vs X residuals vs X

3 $\hat{Y}$ vs X predicted Y vs X