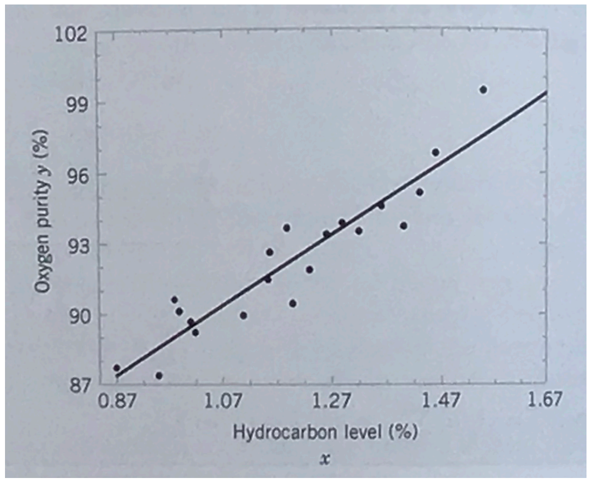
$$\hat{y} = \underset{\hat{\beta}_0}{74.28} + \underset{\hat{\beta}_1}{14.95}\, x$$



- recall the ε term in regression model:

$$Y = 74.28 + 14.95x + \varepsilon$$

error term

difference between expected value @ some x and an actual data point

define residuals as

$$e_i = y_i - \hat{y}_i$$

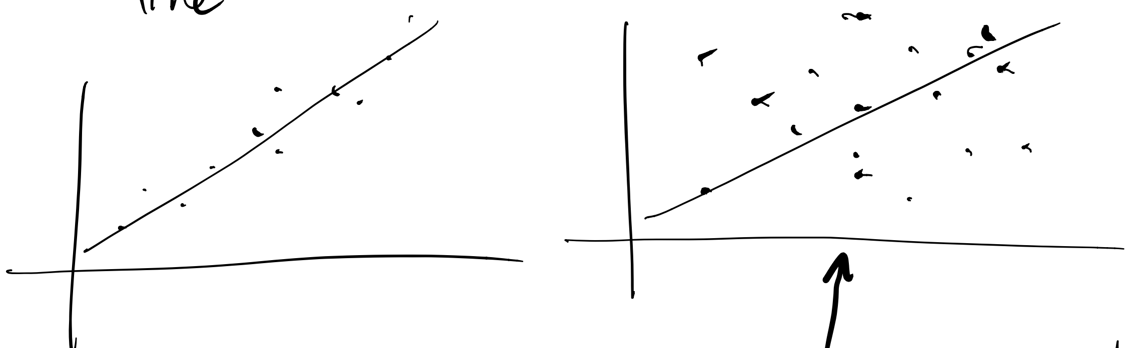
- plotting and analyzing residuals → big deal!

- define <u>error sum of squares</u> (SS_E):

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

- we need to quantify how much "scatter" there is in the data → measure of <u>variance</u>.

- O_2 purity example; "dots" (i.e. data points) are pretty closely grouped around estimated regression line



<u>not</u> the case here!
<u>lots</u> more scatter!

$\therefore$ an unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{SS_E}{n-2}$$

Computational formulae:

$$\underbrace{SS_T}_{\text{total sum of squares}} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \underbrace{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2}_{\text{computational formula}}$$

$$SS_E = SS_T - \hat{\beta}_1 \underbrace{S_{XY}}_{\text{numerator term of } \hat{\beta}_1}$$

Hypothesis Tests in Simple Linear Regression

tests on slope: $H_0 : \beta_1 = \underbrace{\beta_{1,0}}_{\text{some hypothesized value of slope}}$

$H_1 : \beta_1 \neq \beta_{1,0}$

test statistic : $\boxed{T_0 = \dfrac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\hat{\sigma}^2 / S_{xx}}}}$

den. term of $\hat{\beta}_1$

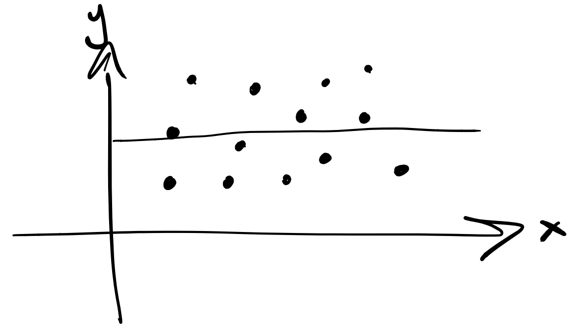
$\boxed{\text{critical values: } \pm t_{\alpha/2,\, n-2}}$

.. important case of hypothesis test on slope :

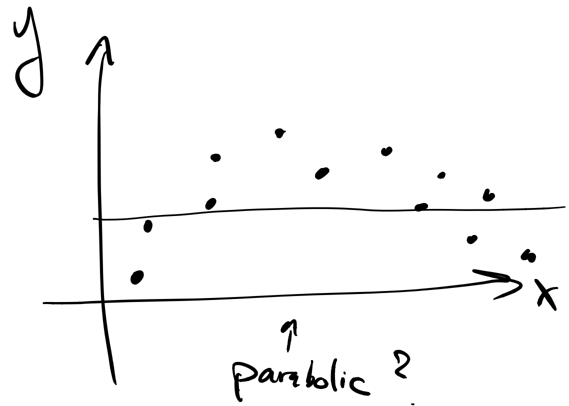
$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

.. failing to reject H_0 could mean one of two things :

1.) there is no significant relationship between Y and X



2.) relationship is not linear



O_2 purity example

test $H_0: \beta_1 = 0$
$H_1: \beta_1 \neq 0$ @ $\alpha = 0.01$

∴ need $\hat{\sigma}^2$ to get T_0

$$SS_T = \sum y_i^2 - n\bar{y}^2 = 170{,}044 - 20 \cdot 92.1605^2$$

$$= 173.4$$

$$SS_E = SS_T - \hat{\beta}_1 S_{xy}$$

$$= 173.4 - 14.947 \cdot 10.17744 = 21.25$$

$$\therefore \hat{\sigma}^2 = \frac{SS_E}{n-2} = \frac{21.25}{18} = 1.181$$

$$t_0 = \frac{\hat{\beta}_1 - \beta_{1,0}^{\to =0}}{\sqrt{\hat{\sigma}^2 / S_{xx}}} = \frac{14.95}{\sqrt{1.181/0.68088}}$$

$$t_0 = 11.35$$!!!!!

Critical values: $\pm t_{\alpha/2, n-2}$

$$= \pm t_{.005, 18} = 2.878$$

$$t_0 >>>>>>>>>>>>> {}^{+}t_{\alpha/2, n-2}$$

strongly reject H_0

Software-generated p-value: 0.0000000123 tests on intercept: $H_0 : \beta_0 = \beta_{0,0}$ hypothesized value of intercept

$H_1 : \beta_0 \neq \beta_{0,0}$

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right]}}$$

O_2 purity example:

$H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$$t_0 = \frac{74.28 - 0}{\sqrt{1.181 \left[\frac{1}{20} + \frac{1.1960^2}{0.68088} \right]}}$$

$$t_0 = 46.61$$

get outta here

(same critical values of $\pm t_{\alpha/2, n-2}$)