Stat 306:
Finding Relationships in Data.
Lecture 6
Section 2.6

Recap from last lecture
2.5 (continued)

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

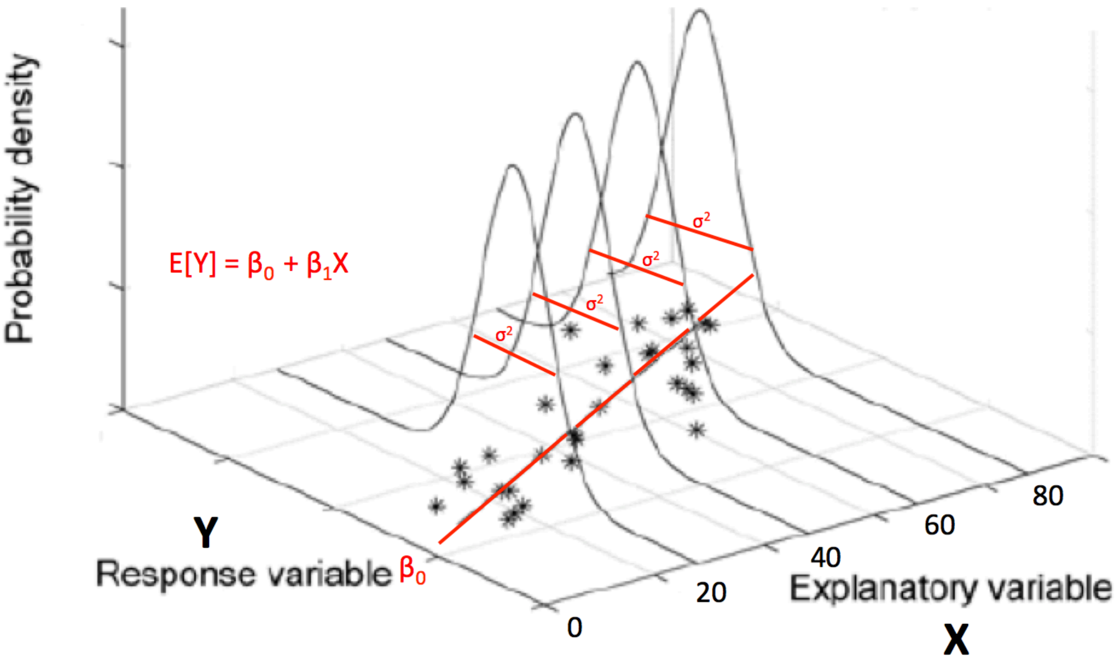
Step 2:
Determine
E[$\hat{\Theta}$] (to confirm it's unbiased)
Var[$\hat{\Theta}$] (to calculate se)

Step 3:
Define
se($\hat{\theta}$) =
estimate of $\sqrt{\mathrm{Var}\,(\hat{\Theta})}$

Step 4:
Define
(1-α)% C.I. =
$\hat{\theta} \pm c \times se(\hat{\theta})$

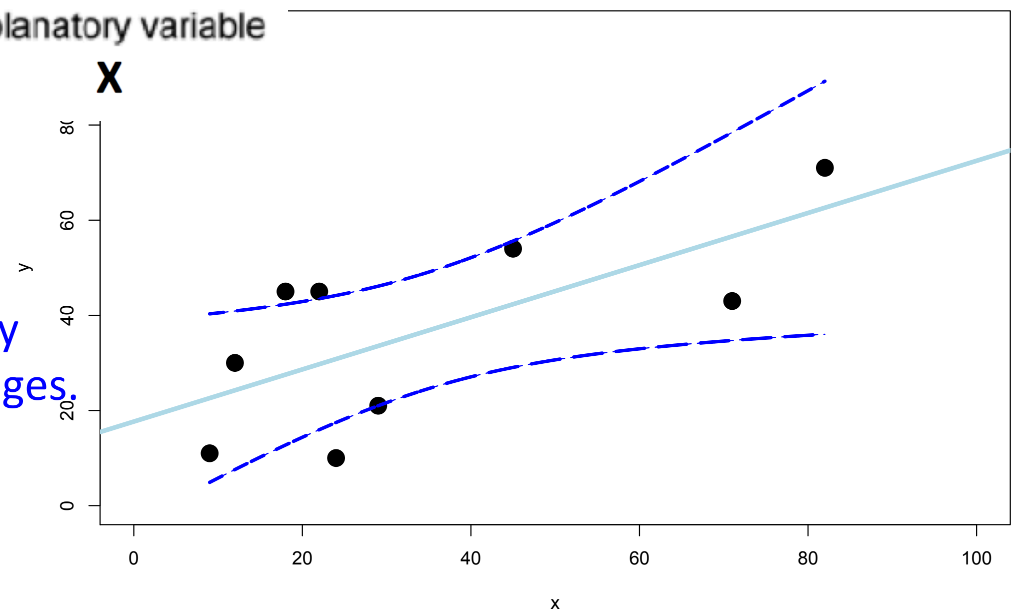
Population parameter or **"something we would like to estimate"**	**Sample statistic ("estimator")**	**Estimator as a Random Variable**	**Expected Value of the estimator**	**Variance of the estimator**	**Standard Error of estimator**	**Confidence Interval**
β_0	b_0	B_0	$E[B_0]$	$Var[B_0]$	$se(b_0)$	C.I. for β_0
β_1	b_1	B_1	$E[B_1]$	$Var[B_1]$	$se(b_1)$	C.I. for β_1
σ^2	s^2	S^2	$E[S^2]$	$Var[S^2]$	$se(s^2)$	C.I. for σ^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$	E $(\hat{\mu}_Y(x))$	Var$(\hat{\mu}_Y(x))$	$se(\hat{\mu}_Y(x))$	C.I. for $\mu_Y(x)$

- **Confused about homogeneity vs. non-consistent width of confidence intervals?**



σ^2 is the variance of Y; constant regardless of the value of x.

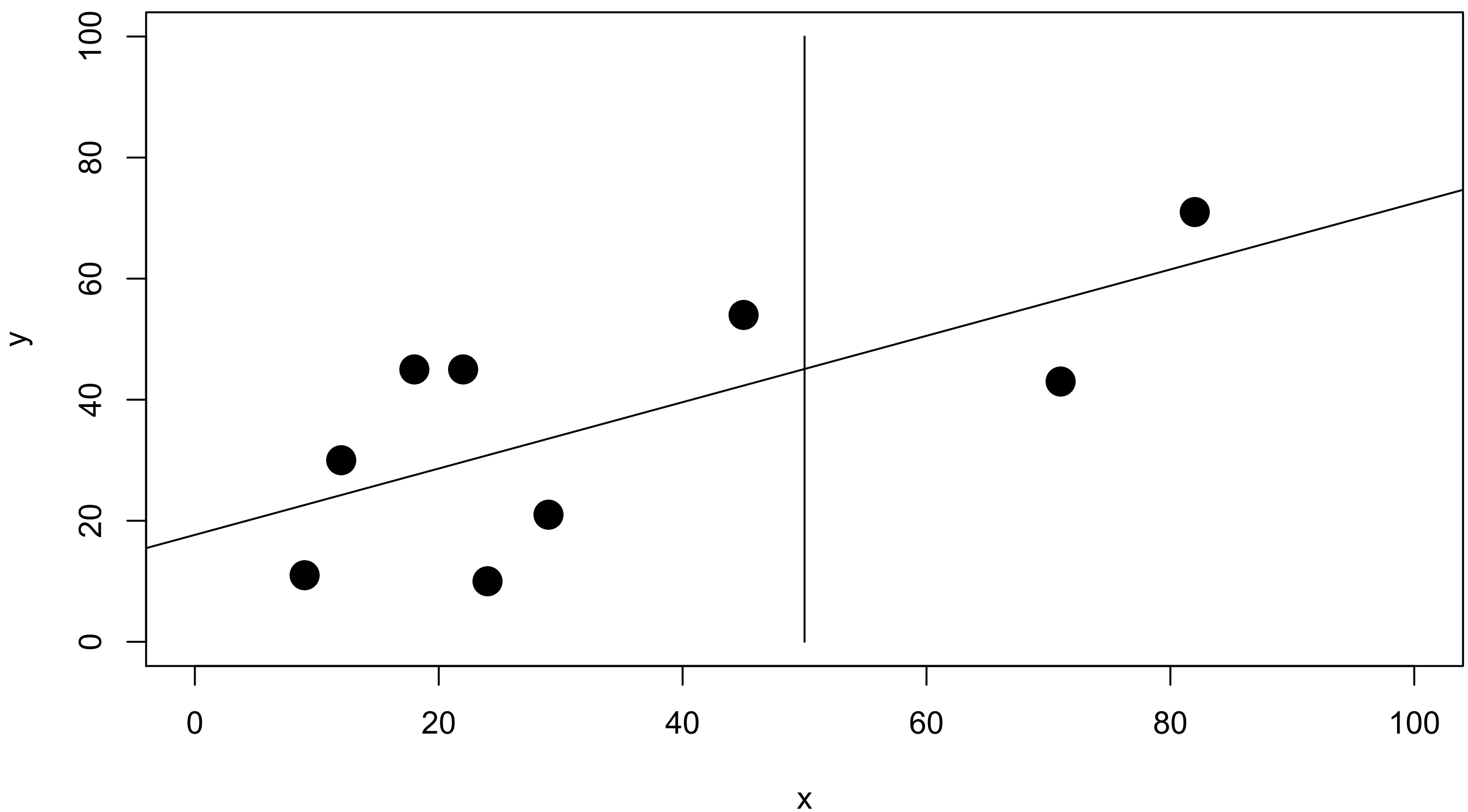
The blue dashed line is the confidence interval for the subpopulation mean.
In other words, it represents the variability in our estimate of the mean of Y as x changes.



Predictions and prediction intervals

Suppose we now want to make a prediction for a new value of x.

Example: Suppose we would like to predict how much money (Y), someone aged 50 years old (X=50) will have.



Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=50 years old will have.

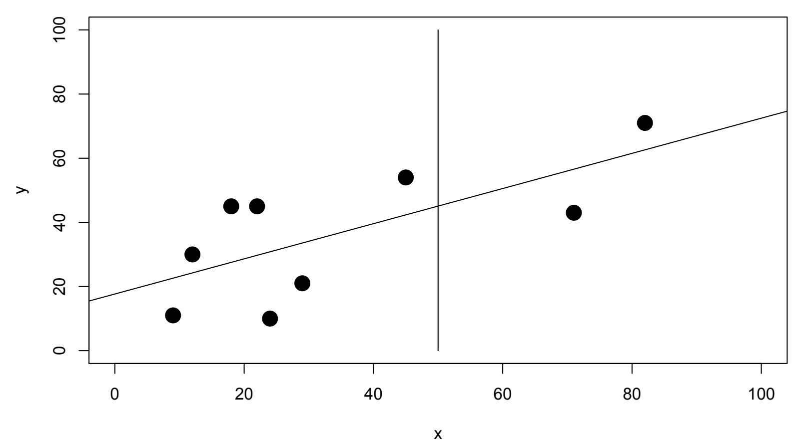
this hypothetical new person aged 50 is sometimes called “an out-of-sample unit with value x*”, Where x*=50.

Our best estimate, also known as the “point prediction”, would be equal to $b_0 + b_1(50) = 45.1$

```
> xstar <- 50
> point_prediction <- beta0hat + beta1hat*xstar
> point_prediction
[1] 45.07117
```

Predictions and prediction intervals

```
> # x and n are fixed values
> x <- c(82, 45, 71, 22, 29, 9, 12, 18, 24)
> n <- 9
>
> # y is a realization of the random variable "Y", i.e. "observed data":
> y <- c(71, 54, 43, 45, 21, 11, 30, 45, 10)
> xbar <- (1/n)*sum(x)
> ybar <- (1/n)*sum(y)
> sx <- sqrt( sum((x-xbar)^2)/(n-1) )
> sy <- sqrt( sum((y-ybar)^2)/(n-1) )
> sxy <- (1/(n-1))*sum((x-xbar)*(y-ybar))
> rxy <- sxy/(sx*sy)
> beta1hat <- rxy*sy/sx
> beta0hat <- ybar-beta1hat*xbar
> residuals <- y - beta0hat - beta1hat*x
> s <- sqrt( (1/(n-2))*sum(residuals^2))
> plot(y~x, xlim=c(0,100), ylim=c(0,100), pch=20, cex=3)
> abline(beta0hat, beta1hat)
>
> xstar <- 50
> point_prediction <- beta0hat + beta1hat*xstar
> point_prediction
[1] 45.07117
> lines(x=c(xstar, xstar),c(0,100))
```



Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=50 years old will have.

$\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)]$$

$$= (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1) x^* - \epsilon(x^*)$$

Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=60 years old will have.

Our prediction

The truth

$\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)]$$
$$= (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*)$$

The difference between our prediction and the truth is the error

Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=60 years old will have.

Our prediction

The truth

$\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)]$$
$$= (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*)$$

The difference between our prediction and the truth is the error

This has variance

$$(2.68) \qquad \mathrm{Var}\,[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] + \mathrm{Var}\,[\epsilon(x^*)] = \sigma^2\left\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\right\} + \sigma^2,$$

Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=60 years old will have.

Our prediction

The truth

$\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67)\quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)]$$

$$= (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*)$$

The difference between our prediction and the truth is the error

Cov() is equal to 0, since the two terms are independent.

This has variance

$$(2.68)\qquad \mathrm{Var}\left[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*\right] + \mathrm{Var}\left[\epsilon(x^*)\right] = \sigma^2\left\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\right\} + \sigma^2,$$

since $\mathrm{Var}\left[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*\right] = \mathrm{Var}\left[\hat{\mu}_Y(x^*)\right]$ from (2.66).

Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=60 years old will have.

Our prediction

The truth

$\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)]$$

$$= (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*)$$

This has variance

$$(2.68) \qquad \mathrm{Var}\,[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] + \mathrm{Var}\,[\epsilon(x^*)] = \sigma^2\left\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\right\} + \sigma^2,$$

So the (estimated) SE of the prediction error is

$$\hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{(n-1)s_x^2}},$$

Note this does not decrease to 0 as $n \to \infty$.

Predictions and prediction intervals

Example: Suppose we would like to predict how much money (Y), someone aged X=60 years old will have.

Our prediction

The truth

$\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)]$$

$$= (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*)$$

This has variance

$$(2.68) \qquad \mathrm{Var}\,[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] + \mathrm{Var}\,[\epsilon(x^*)] = \sigma^2\left\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\right\} + \sigma^2,$$

So the (estimated) SE of the prediction error is

$$\hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{(n-1)s_x^2}};$$

Note this does not decrease to 0 as $n \to \infty$.

Note that variances of estimators include σ^2 in their equations. Estimated SEs replace the "population" quantity σ by a sample quantity $\hat{\sigma}$.

Predictions and prediction intervals

Next for the 95% prediction interval for $Y(x^*)$ for an out-of-sample unit with value x^*, the point prediction is $\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)] = (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*).$$

This has variance

$$(2.68) \qquad \mathrm{Var}\left[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*\right] + \mathrm{Var}\left[\epsilon(x^*)\right] = \sigma^2\left\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\right\} + \sigma^2,$$

since $\mathrm{Var}\left[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*\right] = \mathrm{Var}\left[\hat{\mu}_Y(x^*)\right]$ from (2.66). So the (estimated) SE of the prediction error is

$$(2.69) \qquad \hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{(n-1)s_x^2}},$$

and this does not decrease to 0 as $n \to \infty$.

Note that variances of estimators include σ^2 in their equations. Estimated SEs replace the "population" quantity σ by a sample quantity $\hat{\sigma}$.

Predictions and prediction intervals

Next for the 95% prediction interval for $Y(x^*)$ for an out-of-sample unit with value x^*, the point prediction is $\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

$$(2.67) \quad \hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)] = (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*).$$

This has variance

$$(2.68) \qquad \operatorname{Var}[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] + \operatorname{Var}[\epsilon(x^*)] = \sigma^2\left\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\right\} + \sigma^2,$$

since $\operatorname{Var}[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] = \operatorname{Var}[\hat{\mu}_Y(x^*)]$ from (2.66). So the (estimated) SE of the prediction error is

$$(2.69) \qquad \text{se(E)} = \quad \hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{(n-1)s_x^2}},$$

and this does not decrease to 0 as $n \to \infty$.

Note that variances of estimators include σ^2 in their equations. Estimated SEs replace the "population" quantity σ by a sample quantity $\hat{\sigma}$.

Predictions and prediction intervals

Next for the 95% prediction interval for $Y(x^*)$ for an out-of-sample unit with value x^*, the point prediction is $\hat{Y}(x^*) = \hat{B}_0 + \hat{B}_1 x^*$ with error

(2.67) $\hat{Y}(x^*) - Y(x^*) = \hat{B}_0 + \hat{B}_1 x^* - [\beta_0 + \beta_1 x^* + \epsilon(x^*)] = (\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^* - \epsilon(x^*)$.

This has variance

(2.68) $$\mathrm{Var}\,[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] + \mathrm{Var}\,[\epsilon(x^*)] = \sigma^2\Big\{n^{-1} + \frac{(x^* - \overline{x})^2}{[(n-1)s_x^2]}\Big\} + \sigma^2,$$

since $\mathrm{Var}\,[(\hat{B}_0 - \beta_0) + (\hat{B}_1 - \beta_1)x^*] = \mathrm{Var}\,[\hat{\mu}_Y(x^*)]$ from (2.66). So the (estimated) SE of the prediction error is

(2.69) $$\text{se(E)} = \quad \hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{(n-1)s_x^2}},$$

and this does not decrease to 0 as $n \to \infty$.

Note that variances of estimators include σ^2 in their equations. Estimated SEs replace the "population" quantity σ by a sample quantity $\hat{\sigma}$.

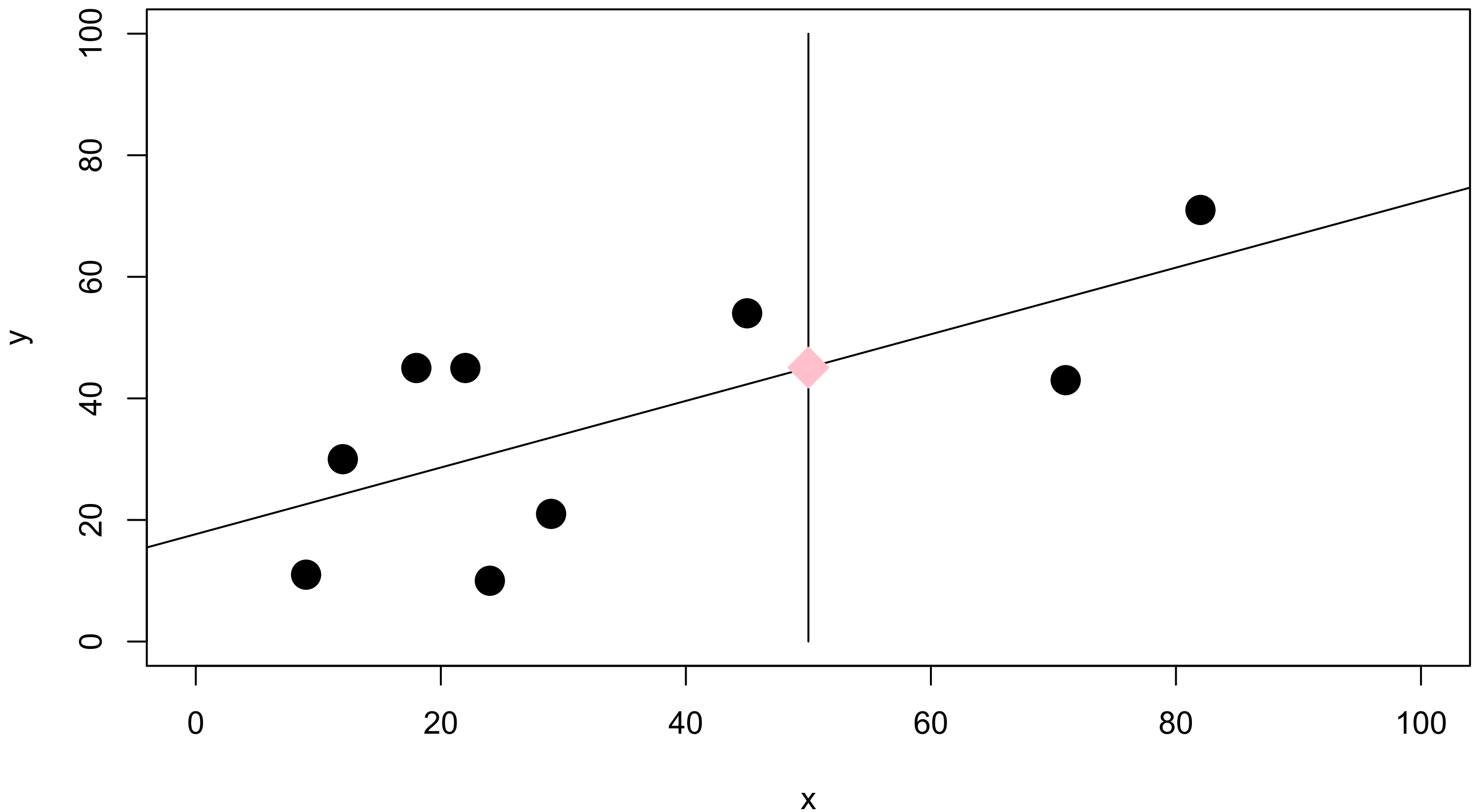
The 95% prediction interval for $Y(x^*)$ for a unit (not in sample) with value x^*:

(2.44) $$\hat{Y}(x^*) \pm t_{n-2,0.975} \times se(E),\ \hat{Y}(x^*) = \hat{\beta}_0 + \hat{\beta}_1 x^* = \hat{\mu}_Y(x^*),$$

where $E = \hat{Y}(x^*) - Y(x^*) = \hat{\mu}_Y(x^*) - Y(x^*) = \hat{\mu}_Y(x^*) - \beta_0 - \beta_1 x^* - \epsilon(x^*)$ is the prediction error.

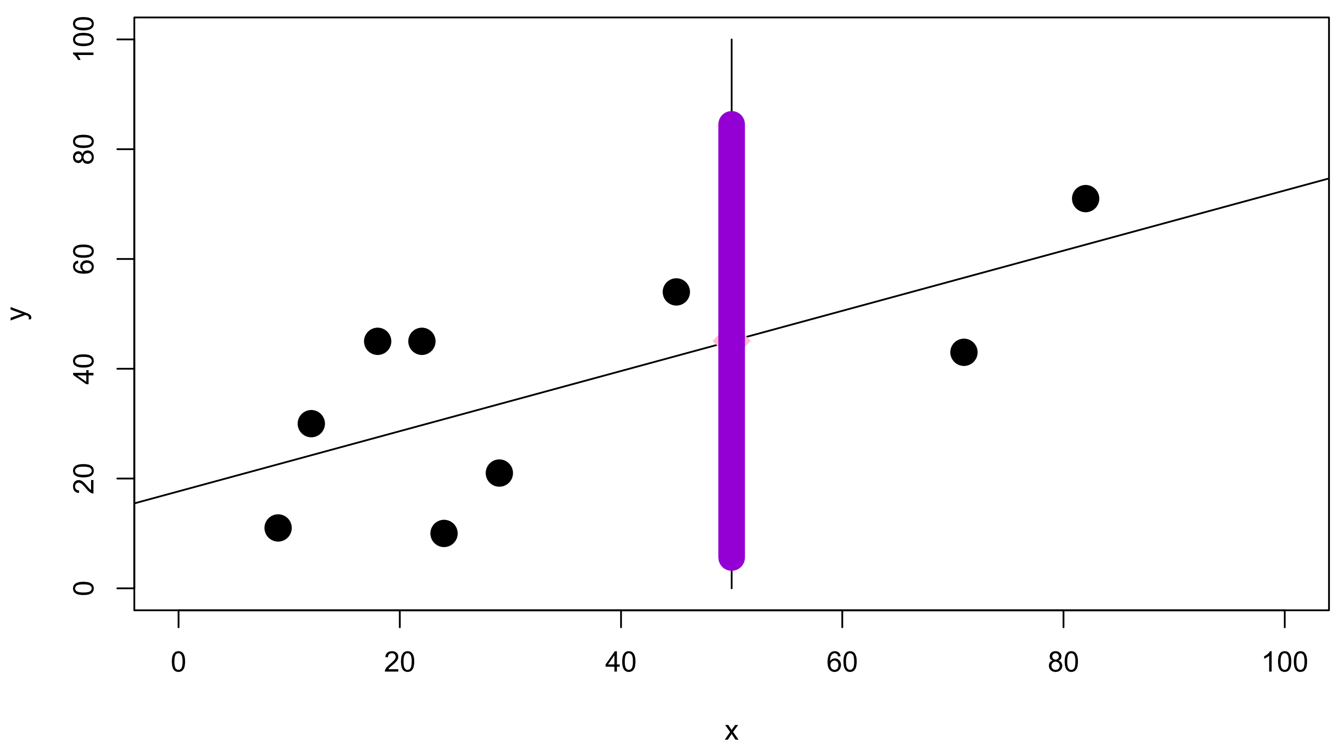
Predictions and prediction intervals

```
> points(xstar, point_prediction, col="pink", pch=18, cex=3)
```



Predictions and prediction intervals

```
> # 95% prediction interval:
> lowerPI <- point_prediction -  qt(0.975,n-2) * s * sqrt(1/n + 1 + ((xstar-xbar)^2)/((n-1)*sx^2))
> upperPI <- point_prediction +  qt(0.975,n-2) * s * sqrt(1/n + 1 + ((xstar-xbar)^2)/((n-1)*sx^2))
>
> c(lowerPI,upperPI)
[1]  5.61226 84.53007
>
> lines(x=c(xstar,xstar),y=c(lowerPI,upperPI), col="darkviolet",lwd=15)
```



Age vs. Money

Sample statistics

$b_0 = 17.7$
$b_1 = 0.55$
$s = 15.5$
$R^2 = 0.49$

For parameter β_1 :

$$95\% \text{ C.I.} = [0.05, 1.05]$$
$$p\text{-value} = 0.036$$

Objective: The purpose of this observational study was to demonstrate if, and to what extent, age is associated with money.

Design and Methods: We collected a random sample of individuals and for each determined their age **(recorded in years)** and the amount of money (in dollars) in their accounts. Analysis of the data was done using **linear regression**.

Results: We obtained a random sample of n = 9 subjects. There is a statistically significant association between age and money (p-value =0.036). For every additional year in age, an individual's amount of money increases on average by an estimated of \$0.55 (95% C.I. = [\$0.05, \$1.05]).

Conclusions: We found that, as hypothesized, age is associated with money. In our sample age accounted for about half of the variability observed in money (R^2=0.49). **We predict that a 50 year old will have \$45.1 (95% P.I. = [\$5.6, \$84.5]),** whereas a 40 year old will have \$39.6 (95% P.I. = [\$0.8, \$78.4]).

Small Print: The analysis rests on the following assumptions:

- the observations are independently and identically distributed.
- the **response** variable, money, is normally distributed.
- Homoscedasticity of residuals or equal variance.
- the relationship between **response** and **predictor** variables is linear.

se(subpopulation mean) VS. se(prediction error)

Subpopulation mean:

$$se(\hat{\mu}_Y(x)) \quad = \quad \hat{\sigma} \times \sqrt{\frac{1}{n} + \frac{(x - \overline{x})^2}{(n-1)s_x^2}}$$

Whereas, the (estimated) SE of the prediction error is:

(2.69) $$\hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{(n-1)s_x^2}},$$

and this does not decrease to 0 as $n \to \infty$.

2.6 Explanation of Student t quantiles in the interval estimates

2.6.1. History lesson about the t-test

2.6.2. Three important things to know about a normal random variable

2.6.3 Estimators as Random Variables (one more time!)

2.6.4 Explanation of Student t quantiles

2.6 Explanation of Student t quantiles in the interval estimates

2.6.1. History lesson about the t-test

2.6.2. Three important things to know about a normal random variable

2.6.3 Estimators as Random Variables (one more time!)

2.6.4 Explanation of Student t quantiles

2.6.1. History lesson about the t-test

Student is the publication pseudonym for William Gosset, who developed methods for inference of means for small samples while working at Guinness Brewery (Ireland) in early 1900s.

https://en.wikipedia.org/wiki/William_Sealy_Gosset

William Sealy Gosset (aka "Student"):

"Is this batch of beer any different than the standard?"

"Let's have a taste test! …t-test anyone?"

2.6 Explanation of Student t quantiles in the interval estimates

2.6.1. History lesson about the t-test

2.6.2. Three important things to know about a normal random variable

2.6.3 Estimators as Random Variables (one more time!)

2.6.4 Explanation of Student t quantiles

2.6.2. Three important things to know about a normal random variable

- **Thing 1:**
  - Linear combinations of independent normal random variables also have normal distributions! (see Appendix B)

2.6.2. Three important things to know about a normal random variable

- **Thing 1:**
  - Linear combinations of independent normal random variables also have normal distributions! (see Appendix B)

For example:

Let:

W_1 be a normal random variable
and W_2 be a normal random variable,

Then:

$W_3 = aW_1 + bW_2$ is a normal r.v.
for any numbers a and b.

2.6.2. Three important things to know about a normal random variable

- **Thing 2:**
  - A normal random variable can be converted to a standard normal random variable.

$$W \sim Normal(\mu, \sigma^2)$$

$$\frac{W-\mu}{\sigma} \sim Normal(0,1)$$

$$Pr(-1.96 < \frac{W-\mu}{\sigma} < 1.96) = 0.95$$

$$Pr(-z_{1-\frac{\alpha}{2}} < \frac{W-\mu}{\sigma} < z_{1-\frac{\alpha}{2}}) = 1-\alpha$$

2.6.2. Three important things to know about a normal random variable

- **Thing 2:**
  - A normal random variable can be converted to a standard normal random variable.

$$W \sim Normal(\mu, \sigma^2)$$

$$Pr(-z_{1-\frac{\alpha}{2}} < \frac{W-\mu}{\sigma} < z_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

$$Pr(W - z_{1-\frac{\alpha}{2}}\sigma < \mu < W + z_{1-\frac{\alpha}{2}}\sigma) = 1 - \alpha$$

For example, with $\alpha = 0.05$:

$$Pr(W - 1.96\sigma < \mu < W + 1.96\sigma) = 0.95$$

2.6.2. Three important things to know about a normal random variable

- **Thing 3:**
  - If the variance is unknown, we must use the t distribution.

$$\frac{W-\mu}{\hat{\sigma}} \sim t_{n-2}$$

$$Pr(-z_{1-\frac{\alpha}{2}} < \frac{W-\mu}{\hat{\sigma}} < z_{1-\frac{\alpha}{2}}) \neq 1-\alpha$$

$$Pr(-t_{n-2,1-\frac{\alpha}{2}} < \frac{W-\mu}{\hat{\sigma}} < t_{n-2,1-\frac{\alpha}{2}}) = 1-\alpha$$

$$Pr(-t_{n-2,1-\frac{\alpha}{2}} < \frac{W-\mu}{\hat{\sigma}} < t_{n-2,1-\frac{\alpha}{2}}) = 0.95$$

for example, with $n = 9$:

$$Pr(-2.26 < \frac{W-\mu}{\hat{\sigma}} < 2.26) = 0.95$$

2.6.2. Three important things to know about a normal random variable

- **Thing 3:**
  - If the variance is unknown, we must use the t distribution.

$$\frac{W-\mu}{\hat{\sigma}} \sim t_{n-2}$$

$$Pr(W - t_{n-2,1-\frac{\alpha}{2}}\hat{\sigma} < \mu < W + t_{n-2,1-\frac{\alpha}{2}}\hat{\sigma}) = 1 - \alpha$$

for example, with $n = 9$:
$$Pr(W - 2.26\hat{\sigma} < \mu < W + 2.26\hat{\sigma}) = 0.95$$

2.6 Explanation of Student t quantiles in the interval estimates

2.6.1. History lesson about the t-test

2.6.2. Three important things to know about a normal random variable

2.6.3 Estimators as Random Variables (one more time!)

2.6.4 Explanation of Student t quantiles

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Step 2:
Determine
E[$\hat{\Theta}$] (to confirm it's unbiased)
Var[$\hat{\Theta}$] (to calculate se)

Step 3:
Define
se($\hat{\theta}$) =
estimate of $\sqrt{\mathrm{Var}(\hat{\Theta})}$

Step 4:
Define
(1-α)% C.I. =
$\hat{\theta} \pm c \times se(\hat{\theta})$

Population parameter or "something we would like to estimate"	**Sample statistic ("estimator")**	**Estimator as a Random Variable**	**Expected Value of the estimator**	**Variance of the estimator**	**Standard Error of estimator**	**Confidence Interval**
β_0	b_0	B_0	$E[B_0]$	$Var[B_0]$	$se(b_0)$	C.I. for β_0
β_1	b_1	B_1	$E[B_1]$	$Var[B_1]$	$se(b_1)$	C.I. for β_1
σ^2	s^2	S^2	$E[S^2]$	$Var[S^2]$	$se(s^2)$	C.I. for σ^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$	$E(\hat{\mu}_Y(x))$	$Var(\hat{\mu}_Y(x))$	$se(\hat{\mu}_Y(x))$	C.I. for $\mu_Y(x)$

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Population parameter or "something we would like to estimate"	**Sample statistic ("estimator")**	**Estimator as a Random Variable**
β_0	b_0	B_0
β_1	b_1	B_1
σ^2	s^2	S^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$

Recall:

$$B_1 = \sum_{i=1}^{n} a_i Y_i$$

, where: $a_i = \frac{(x_i - \bar{x})}{(n-1)s_x^2}$

and:

$$\mathrm{E}(\hat{B}_1) = \sum^{n} a_i \mathrm{E}(Y_i) = \sum_{i=1}^{n} a_i(\beta_0 + \beta_1 x_i) = \beta_1,$$

and:

$$\mathrm{Var}(\hat{B}_1) = \frac{\sigma^2}{(n-1)s_x^2}$$

Since B_1 is a linear combination of the Y_is (Normal RVs), then **(with Thing 1)**:

$$\hat{B}_1 \sim N\left(\beta_1, \sigma^2_{\hat{\beta}_1} = \frac{\sigma^2}{(n-1)s_x^2}\right),$$

$$\frac{\hat{B}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim N(0,1).$$

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Population parameter or "something we would like to estimate"	Sample statistic ("estimator")	Estimator as a Random Variable
β_0	b_0	B_0
β_1	b_1	B_1
σ^2	s^2	S^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$

Recall:

$$
\begin{aligned}
E[B_0] &= E[\bar{Y} - B_1\bar{X}] \\
&= \tfrac{1}{n}E[\textstyle\sum_{i=1}^{n} Y_i] - \beta_1\bar{X} \\
&= \tfrac{1}{n}\textstyle\sum_{i=1}^{n}(\beta_0 + \beta_1 X_i) - \beta_1\bar{X} \\
&= \beta_0 + \tfrac{1}{n}\textstyle\sum_{i=1}^{n}\beta_1 X_i - \beta_1\bar{X} \\
&= \beta_0 + \beta_1\tfrac{1}{n}\textstyle\sum_{i=1}^{n} X_i - \beta_1\bar{X} \\
&= \beta_0 + \beta_1\bar{X} - \beta_1\bar{X} \\
&= \beta_0
\end{aligned}
$$

Also:

$$
\begin{aligned}
Var(B_0) &= Var(\textstyle\sum_{i=1}^{n} Y_i/n) + \bar{X}^2 Var(B_1) \\
&= \sigma^2\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right)
\end{aligned}
$$

For the intercept, we can, again, make use of the fact that B_0 is a linear combination of normal random variables **(Thing 1)**:

$$B_0 \sim N\left[\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)\right]$$

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Population parameter or "something we would like to estimate"	**Sample statistic ("estimator")**	**Estimator as a Random Variable**
β_0	b_0	B_0
β_1	b_1	B_1
σ^2	s^2	S^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$

$$S^2/\sigma^2 \sim \chi^2_{n-2}$$

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Population parameter or "something we would like to estimate"	**Sample statistic ("estimator")**	**Estimator as a Random Variable**
β_0	b_0	B_0
β_1	b_1	B_1
σ^2	s^2	S^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$

We have that:

$$\mathrm{E}\left[\hat{\mu}_Y(x)\right] = \beta_0 + \beta_1 x$$

$$\mathrm{Var}\left[\hat{\mu}_Y(x)\right] = \sigma^2\left\{n^{-1} + \frac{(x-\overline{x})^2}{[(n-1)s_x^2]}\right\}$$

And again, a linear combination of normal random variables is a normal random variable **(Thing 1)**:

$$\mu_Y(x) \sim Normal\left(\beta_0 + \beta_1 x, \sigma^2\left(\frac{1}{n} + \frac{(x-\bar{x})^2}{[(n-1)s_x^2]}\right)\right)$$

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Population parameter **or "something we would like to estimate"**	**Sample statistic ("estimator")**	**Estimator as a Random Variable**	
β_0	b_0	B_0	$B_0 \sim N[\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)]$
β_1	b_1	B_1	$\hat{B}_1 \sim N\left(\beta_1, \sigma^2_{\hat{\beta}_1} = \frac{\sigma^2}{(n-1)s_x^2}\right)$
σ^2	s^2	S^2	$S^2/\sigma^2 \sim \chi^2_{n-2}$
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$	$\mu_Y(x) \sim Normal\left(\beta_0 + \beta_1 x, \sigma^2(\frac{1}{n} + \frac{(x-\bar{x})^2}{[(n-1)s_x^2]})\right)$

2.6 Explanation of Student t quantiles in the interval estimates

2.6.1. History lesson about the t-test

2.6.2. Three important things to know about a normal random variable

2.6.3 Estimators as Random Variables (one more time!)

2.6.4 Explanation of Student t quantiles

Step 0:
From θ, define estimator, $\hat{\theta}$

Step 1:
Consider the sample statistic, $\hat{\theta}$, as a random variable $\hat{\Theta}$

Population parameter or "something we would like to estimate"	**Sample statistic ("estimator")**	**Estimator as a Random Variable**
β_0	b_0	B_0
β_1	b_1	B_1
σ^2	s^2	S^2
$\mu_Y(x)$	$(\hat{\mu}_Y(x))$	$(\hat{\mu}_Y(x))$

$$B_0 \sim N[\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)]$$

With **Thing 2**, we have:

$$\frac{B_0-\beta_0}{\sqrt{\sigma^2\left(\frac{1}{n}+\frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)}} \sim N(0,1)$$

But we do not know the variance. We only have an estimate of the variance, so (with **Thing 3**):

$$\frac{B_0-\beta_0}{SE(B_0)} \sim t_{n-2}$$

And therefore:

95% C.I. for β_0 =

$$\left[\ \bar{y} - b_1\bar{X} - t_{n-2,0.975}\cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right)}\ ,\ \bar{y} - b_1\bar{X} + t_{n-2,0.975}\cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right)}\ \right]$$

2.6.4 Explanation of Student t quantiles

Estimator as a Random Variable
B_0
B_1
S^2
$(\hat{\mu}_Y(x))$

With **Thing 1**, we have:

$$\hat{B}_1 \sim N\left(\beta_1, \sigma^2_{\hat{\beta}_1} = \frac{\sigma^2}{(n-1)s_x^2}\right),$$

With **Thing 2**, we have:

$$\frac{\hat{B}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim N(0,1).$$

But we do not know the variance,
so with **Thing 3**, we have:

$$\frac{B_1 - \beta_1}{SE(B_1)} \sim t_{n-2}$$

And therefore:

$$Pr(b_1 - t_{n-2,1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{(n-1)}s_x} < \beta_1 < b_1 + t_{n-2,1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{(n-1)}s_x}) = 1 - \alpha$$

2.6.4 Explanation of Student t quantiles

Estimator as a Random Variable
B_0
B_1
S^2
$(\hat{\mu}_Y(x))$

With **Thing 1**, we have:

$$\hat{B}_1 \sim N\left(\beta_1, \sigma^2_{\hat{\beta}_1} = \frac{\sigma^2}{(n-1)s_x^2}\right),$$

With **Thing 2**, we have:

$$\frac{\hat{B}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim N(0,1).$$

But we do not know the variance,
so with **Thing 3**, we have:

$$\frac{B_1 - \beta_1}{SE(B_1)} \sim t_{n-2}$$

And therefore:

$$Pr(b_1 - t_{n-2,1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{(n-1)}s_x} < \beta_1 < b_1 + t_{n-2,1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{(n-1)}s_x}) = 1 - \alpha$$

with α=0.05:

$$Pr(b_1 - t_{n-2,0.975} \frac{\hat{\sigma}}{\sqrt{(n-1)}s_x} < \beta_1 < b_1 + t_{n-2,0.975} \frac{\hat{\sigma}}{\sqrt{(n-1)}s_x}) = 0.95$$

2.6.4 Explanation of Student t quantiles

Estimator as a Random Variable
B_0
B_1
S^2
$(\hat{\mu}_Y(x))$

With **Thing 1**, we have:

$$\hat{B}_1 \sim N\left(\beta_1, \sigma^2_{\hat{\beta}_1} = \frac{\sigma^2}{(n-1)s_x^2}\right),$$

With **Thing 2**, we have:

$$\frac{\hat{B}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim N(0,1).$$

But we do not know the variance,
so with **Thing 3**, we have:

$$\frac{B_1 - \beta_1}{SE(B_1)} \sim t_{n-2}$$

And therefore:

95% C.I. for β_1 = $\left[b_1 - t_{n-2,0.975}\frac{\hat{\sigma}}{\sqrt{n-1}s_x}, \quad b_1 + t_{n-2,0.975}\frac{\hat{\sigma}}{\sqrt{n-1}s_x}\right]$

where: $b_1 = r_{xy}\frac{s_y}{s_x}$, $\hat{\sigma} = s = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-2}}$

2.6.4 Explanation of Student t quantiles

Estimator as a Random Variable
B_0
B_1
S^2
$(\hat{\mu}_Y(x))$

With **Thing 1**, we have:

$$B_0 \sim N[\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)]$$

With **Thing 2**, we have:

$$\frac{B_0 - \beta_0}{\sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)} \sim N[0, 1]$$

But we do not know the variance, so with **Thing 3**, we have:

$$\frac{B_0 - \beta_0}{SE(B_0)} \sim t_{n-2}$$

And therefore:

$$Pr\left(b_0 - t_{n-2,1-\frac{\alpha}{2}} \cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)} < \beta_0 < b_0 + t_{n-2,1-\frac{\alpha}{2}} \cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}\right) = 1 - \alpha$$

with α=0.05:

$$Pr\left(b_0 - t_{n-2,0.975} \cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)} < \beta_0 < b_0 + t_{n-2,0.975} \cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}\right) = 0.95$$

2.6.4 Explanation of Student t quantiles

Estimator as a Random Variable
B_0
B_1
S^2
$(\hat{\mu}_Y(x))$

With **Thing 1**, we have:

$$B_0 \sim N[\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)]$$

With **Thing 2**, we have:

$$\frac{B_0 - \beta_0}{\sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)} \sim N[0, 1]$$

But we do not know the variance, so with **Thing 3**, we have:

$$\frac{B_0 - \beta_0}{SE(B_0)} \sim t_{n-2}$$

And therefore:

$$95\% \text{ C.I. for } \beta_0 =$$

$$\left[\quad \bar{y} - b_1\bar{x} - t_{n-2,0.975} \cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)} \quad , \quad \bar{y} - b_1\bar{x} + t_{n-2,0.975} \cdot s\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)} \quad \right]$$

2.6.4 Explanation of Student t quantiles

Estimator as a Random Variable
B_0
B_1
S^2
$(\hat{\mu}_Y(x))$

With **Thing 1**, we have:

$$\mu_Y(x) \sim Normal\left(\beta_0 + \beta_1 x, \sigma^2\left(\frac{1}{n} + \frac{(x-\bar{x})^2}{[(n-1)s_x^2]}\right)\right)$$

With **Thing 2**, we have…

But we do not know the variance,
so with **Thing 3**, we have…

And therefore:

The 95% confidence interval for subpopulation mean $\mu_Y(x) = \beta_0 + \beta_1 x$ is

$$\hat{\mu}_Y(x) \pm t_{n-2,0.975} \times se(\hat{\mu}_Y(x)),$$

where:

$$se(\hat{\mu}_Y(x)) = \hat{\sigma} \times \sqrt{\frac{1}{n} + \frac{(x-\overline{x})^2}{(n-1)s_x^2}}$$

$$\hat{\mu}_Y(x) = \hat{\beta}_0 + \hat{\beta}_1 x.$$

$$\hat{\sigma} = s = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-2}}$$

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

Hypothesis Test

$H_0 : \beta_1 = 0$ "Null" hypothesis

$H_1 : \beta_1 \neq 0$

"Alternative" hypothesis

We have:

$$\frac{B_1 - \beta_1}{SE(B_1)} \sim t_{n-2}$$

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

Hypothesis Test

$$H_0 : \beta_1 = 0$$

"Null" hypothesis

$$H_1 : \beta_1 \neq 0$$

"Alternative" hypothesis

We have:

$$\frac{B_1 - \overbrace{\cancel{\beta_1}}^{= 0}}{SE(B_1)} \sim t_{n-2}$$

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

Hypothesis Test

$$H_0 : \beta_1 = 0$$

"Null" hypothesis

$$H_1 : \beta_1 \neq 0$$

"Alternative" hypothesis

We have:

$$\frac{B_1 - \underbrace{\cancel{\beta_1}}_{=0}}{SE(B_1)} \sim t_{n-2}$$

Therefore, "under the null", we have:

$$\frac{B_1}{SE(B_1)} \sim t_{n-2}$$

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

Even if we decide to record "Age" (x) in months and "Money" (Y) in pennies, "under the null", we still have:

$$\frac{B_1}{SE(B_1)} \sim t_{n-2}$$

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

Even if we decide to record "Age" (x) in months and "Money" (Y) in pennies, "under the null", we still have:

$$\frac{B_1}{SE(B_1)} \sim t_{n-2}$$

and:

$$t\text{-statistic} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$$

Therefore…

If β_1 (the slope) was actually equal to 0, it would be very unlikely that the absolute t-stat would be very large.

```
> x <- c(82, 45, 71, 22, 29, 9, 12, 18, 24)
> y <- c(71, 54, 43, 45, 21, 11, 30, 45, 10)
> n <- 9
> xbar <- (1/n)*sum(x)
> sx <- sqrt( sum((x-xbar)^2)/(n-1) )
> ybar <- (1/n)*sum(y)
> sy <- sqrt( sum((y-ybar)^2)/(n-1) )
> sxy <- (1/(n-1))*sum((x-xbar)*(y-ybar))
> rxy <- sxy/(sx*sy)
> b0_hat<-ybar-b1_hat*xbar
> b1_hat <- rxy*sy/sx
> residuals <- y - b0_hat - b1_hat*x
> s <- sqrt( (1/(n-2))*sum(residuals^2) )
> s
[1] 15.5308
> SE_b1 <- s/(sqrt(n-1)*sx)
> SE_b1
[1] 0.2108794
> tstat_b1 <- b1_hat/SE_b1
> tstat_b1
[1] 2.599209
```

2.6.4 Explanation of Student t quantiles

For the null hypothesis $H_0 : \beta_1 = 0$. (2.76) implies that the null distribution of $\hat{B}_1/SE(\hat{B}_1)$ is t_{n-2}. For the data version, $\hat{\beta}_1/se(\hat{\beta}_1)$ is the standardized version of $\hat{\beta}_1$; it is invariant to scale changes of the x and y variables (because a scale change affect the SE in the same way as $\hat{\beta}_1$). $|\hat{\beta}_1/se(\hat{\beta}_1)|$ is the absolute t-ratio statistic and large values indicate that the slope is significantly different from 0.

Even if we decide to record "Age" (x) in months and "Money" (Y) in pennies, "under the null", we still have:

$$\frac{B_1}{SE(B_1)} \sim t_{n-2}$$

and:

$$t\text{-statistic} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$$

Therefore…

If β_1 (the slope) was actually equal to 0, it would be very unlikely that the absolute t-stat would be very large.

```
> b1_hat <- rxy*sy/sx
> b1_hat
[1] 0.5481195
> residuals <- y - b0_hat - b1_hat*x
> s <- sqrt( (1/(n-2))*sum(residuals^2) )
> s
[1] 15.5308
> SE_b1 <- s/(sqrt(n-1)*sx)
> SE_b1
[1] 0.2108794
> tstat_b1 <- b1_hat/SE_b1
> tstat_b1
[1] 2.599209
>
> # How unlikely would it be to obtain
> # a value this large (or larger) if the
> # null was true (i.e. beta_1 = 0)?
>
> 2*pt(abs(tstat_b1), n-2, lower=FALSE)
[1] 0.03546599
```

2.6.4 Explanation of Student t quantiles

```
> summary(lm(y~x))

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-20.820 -12.561   5.757  11.669  17.469 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)  17.6652     8.9579   1.972   0.0892 .
x             0.5481     0.2109   2.599   0.0355 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.53 on 7 degrees of freedom
Multiple R-squared:  0.4911,	Adjusted R-squared:  0.4184 
F-statistic: 6.756 on 1 and 7 DF,  p-value: 0.03547
```

```
> b1_hat <- rxy*sy/sx
> b1_hat
[1] 0.5481195
> residuals <- y - b0_hat - b1_hat*x
> s <- sqrt( (1/(n-2))*sum(residuals^2) )
> s
[1] 15.5308
> SE_b1 <- s/(sqrt(n-1)*sx)
> SE_b1
[1] 0.2108794
> tstat_b1 <- b1_hat/SE_b1
> tstat_b1
[1] 2.599209
>
> # How unlikely would it be to obtain
> # a value this large (or larger) if the
> # null was true (i.e. beta_1 = 0)?
>
> 2*pt(abs(tstat_b1), n-2, lower=FALSE)
[1] 0.03546599
```

2.6.4 Explanation of Student t quantiles

```
> b0_hat<-ybar-b1_hat*xbar
> b0_hat
[1] 17.66519
> residuals <- y - b0_hat - b1_hat*x
> s <- sqrt( (1/(n-2))*sum(residuals^2) )
> s
[1] 15.5308
> SE_b0 <- s*sqrt( 1/n + ( (xbar^2)/( sum((x-xbar)^2)) ) )
> SE_b0
[1] 8.957892
> tstat_b0 <- b0_hat/SE_b0
> tstat_b0
[1] 1.972026
>
> # How unlikely would it be to obtain
> # a value this large (or larger) if the
> # null was true (i.e. beta_0 = 0)?
>
> 2*pt(abs(tstat_b0), n-2, lower=FALSE)
[1] 0.08922444
```

```
> summary(lm(y~x))

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-20.820 -12.561   5.757  11.669  17.469

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.6652     8.9579   1.972   0.0892 .
x             0.5481     0.2109   2.599   0.0355 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.53 on 7 degrees of freedom
Multiple R-squared:  0.4911, Adjusted R-squared:  0.4184
F-statistic: 6.756 on 1 and 7 DF,  p-value: 0.03547
```

2.6.4 Explanation of Student t quantiles

Interpretation for a confidence interval: be careful in what is correct and incorrect usage.
A confidence interval consists of an interval estimate of a population parameter (Greek letter). So we can say 95% confidence interval for β_1 but **not** 95% confidence interval for $\hat{\beta}_1$.

2.6.4 Explanation of Student t quantiles

Interpretation for a confidence interval: be careful in what is correct and incorrect usage.
A confidence interval consists of an interval estimate of a population parameter (Greek letter). So we can say 95% confidence interval for β_1 but **not** 95% confidence interval for $\hat{\beta}_1$.
The interval (2.79) has 95% probability content if $\hat{B}_1$ and $\hat{S} = \hat{\sigma}$ are considered as random variables.
When $\hat{\beta}_1$ and $\hat{\sigma}$ are computed from data values, the interval (2.79) is called a 95% confidence interval.

2.6.4 Explanation of Student t quantiles

Interpretation for a confidence interval: be careful in what is correct and incorrect usage.
A confidence interval consists of an interval estimate of a population parameter (Greek letter). So we can say 95% confidence interval for β_1 but **not** 95% confidence interval for $\hat{\beta}_1$.
The interval (2.79) has 95% probability content if $\hat{B}_1$ and $\hat{S} = \hat{\sigma}$ are considered as random variables.
When $\hat{\beta}_1$ and $\hat{\sigma}$ are computed from data values, the interval (2.79) is called a 95% confidence interval.
For example, the 95% confidence interval for the Merck beta is $0.871 \pm 0.497 = (0.374, 1.368)$.
With numbers (not random variables) in the interval, the interval either contains the true β_1 or it doesn't (and probability is 1 or 0). This is the reason why a numerical interval of most plausible values for a population parameter is called a confidence interval. Probability content of an interval containing a quantity can only be considered if the endpoints of the interval are considered as random variables and not a specific numbers computed from data.

2.6.4 Explanation of Student t quantiles

Explanation of **Thing 3**:

Definition of a t_ν random variable.

Let $Z \sim N(0,1)$ and let $W \sim \chi^2_\nu$ (chi-square distribution with ν degrees of freedom, this has a right-skewed density on the positive real line). The random variables Z and W are mutually independent. Then the definition of a t_ν random variables from a standard normal random variable and a chi-square random variable is as follows:

$$T = \frac{Z}{\sqrt{W/\nu}} \sim t_\nu$$

or T has a Student t_ν distribution. For this to apply above to $\frac{\hat{B}_1 - \beta_1}{SE(\hat{B}_1)}$ in (2.76), write

$$\frac{\hat{B}_1 - \beta_1}{SE(\hat{B}_1)} = \frac{(\hat{B}_1 - \beta_1)/\sigma_{\hat{\beta}_1}}{\sqrt{SE^2(\hat{B}_1)/\sigma^2_{\hat{\beta}_1}}}.$$

Since $Z = (\hat{B}_1 - \beta_1)/\sigma_{\hat{\beta}_1} \sim N(0,1)$, to get a t_{n-2} distribution, it is necessary to show that $SE^2(\hat{B}_1)$ (as a random variable) is independent of $\hat{B}_1$, and that $W = (n-2)SE^2(\hat{B}_1)/\sigma^2_{\hat{\beta}_1} \sim \chi^2_{n-2}$.

linear regression

Age vs. Money

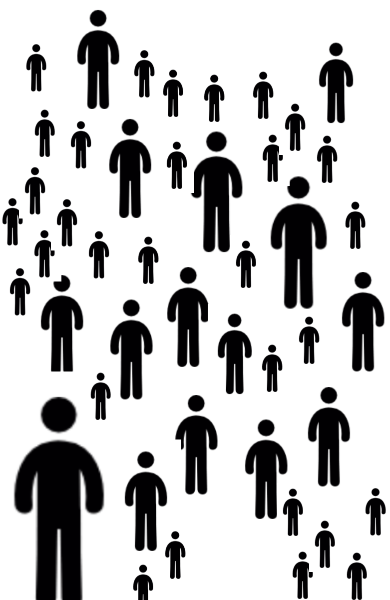
PREDICTOR *variable*

$x \longrightarrow$ Age in Years

RESPONSE *variable*

$Y \longrightarrow$ dollars ($) In bank account

Population



Population parameters

$$\beta_0, \beta_1, \sigma^2$$

Hypothesis Test

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

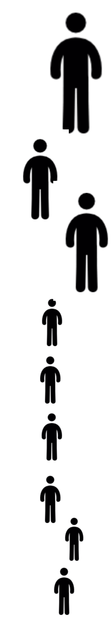
Sample statistics

b_0 = 17.7
b_1 = 0.55
s = 15.5
R^2 = 0.49

For parameter β_1 :

95% C.I. = [0.05, 1.05]
p-value = 0.036

Sample, n=9



x	y
82	71
45	54
71	43
22	45
29	21
9	11
12	30
18	45
24	10