

SEMINAR 1

2.4

X is Bernoulli (uniform(0,1) discrete).

$$P(X=1) = p \quad (\Rightarrow P(X=0) = 1-p)$$

We know that by definition (in general)

$$E(Y) = y_1 p_1 + y_2 p_2 \quad \text{when } Y \text{ takes two values}$$

or more generally $E(Y) = \sum_{i=1}^k y_i p_i$

when Y takes k values.

Hence in our case;

a) $E(X^3) = 1^3 \times p + 0^3 \times (1-p) = p$

b) $E(X^k) = 1^k \times p = 0^k \times (1-p) = p$

b) Mean is simply a different word for expectation.

$$\text{Hence } E(X) = p = \underline{0,3}$$

Variance of X has the formula

$$\text{var}(X) = E(X^2) - (E(X))^2$$

$$\text{var}(X) = p - p^2 = p(1-p) = 0,3 \times 0,7 = \underline{0,21}$$

Skewness is a measure of the deviation from symmetry; where 0 is no asymmetry. So the intuition is that we measure the extent to which the distribution "stems" away from its mean to either side.

$$\text{Skewness} = \frac{E[(X - E(X))^3]}{\sigma_X^3}$$

σ_X^3 is simply the standard deviation cubed.

The standard deviation is the square-root of the variance.

$$\text{st. dev}(X) = \sqrt{\text{var}(X)} = \sqrt{p(1-p)} = \sqrt{0,21} \approx 0,4583$$

○ Using the formula

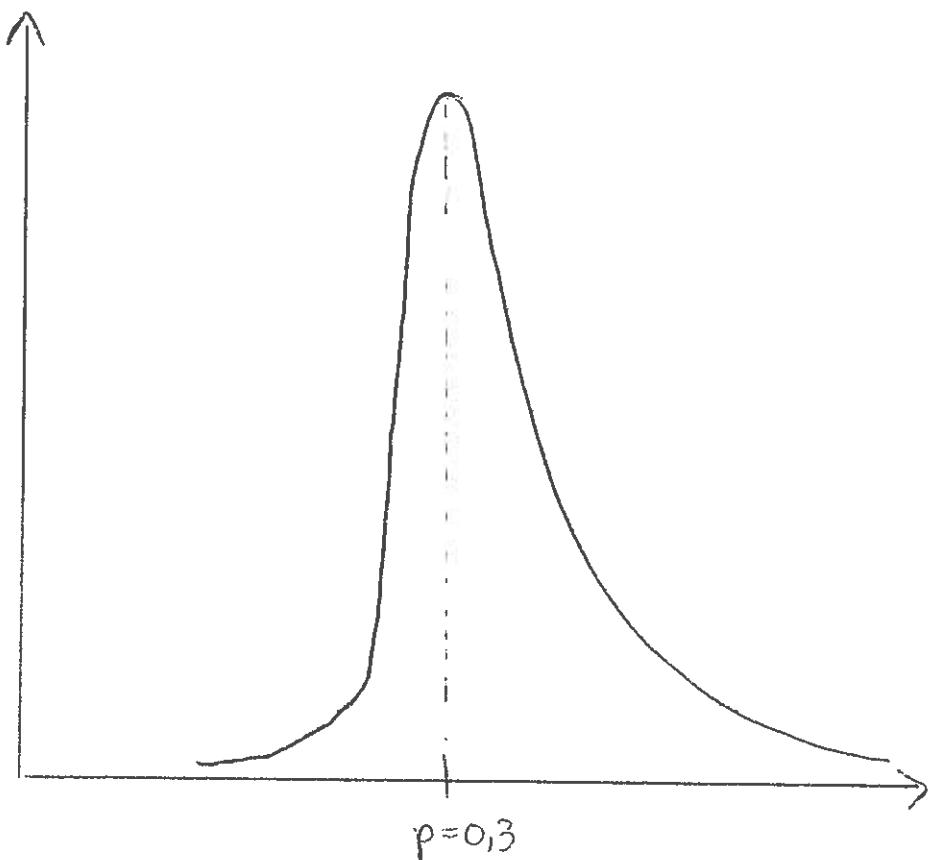
$$E(X-\mu)^3 = E(X^3) - 3[E(X^2)][E(X)] + 2[E(X)]^3$$

$$= p - 3p^2 + 2p^3 = 0,3 - 3 \times 0,3^2 + 2 \times 0,3^3$$

$$= 0,084$$

○ We get Skewness = $\frac{0,084}{0,4583} \approx \underline{\underline{0,183}} 0,87$

which graphically means we have a minor skewness to the left



Kurtosis is a measure of the "fatness" of the tail in a distribution.

$$\text{Kurtosis} = \frac{E[(X - E(X))^4]}{\sigma_x^4}$$

σ_x^4 is simply the squared variance.

$$\Rightarrow \sigma_x^4 = 0,21^2 = 0,0441$$

Using the formula

$$\begin{aligned} E(X - \mu)^4 &= E(X^4) - 4[E(X)][E(X^3)] + 6[E(X)]^2[E(X^2)] \\ &\quad - 3[E(X)]^4 \\ &= p - 4p^2 + 6p^3 - 3p^4 \approx 0,0777 \end{aligned}$$

$$\Rightarrow \text{Kurtosis} = \frac{0,0777}{0,0441} \approx \underline{\underline{1,7619}}$$

① 2.8

$$Y \sim (1, 4) \quad (\mu, \sigma^2)$$

$$Z = \frac{1}{2}(Y-1)$$

$$E(Z) = E\left(\frac{1}{2}(Y-1)\right) = \frac{1}{2}E(Y-1) = \frac{1}{2}[E(Y)-E(1)]$$

$$E(Z) = \frac{1}{2}[E(Y)-1] = \frac{1}{2}[1-1] = \underline{\underline{0}}$$

$$\text{var}(Z) = \text{var}\left(\frac{1}{2}(Y-1)\right) = \frac{1}{4}\text{var}(Y-1)$$

$$\text{var}(Z) = \frac{1}{4}\text{var}(Y) = \frac{1}{4} \times 4 = \underline{\underline{1}}$$

2.17

$Y_i \quad i=1, 2, \dots, n \sim \text{iid } (0,1) \text{ Bernoulli}$

$$p = 0.4$$

$$\text{Then} - \sigma^2_Y = 0.4(1-0.4) = \underline{0.24}$$

By the central limit theorem, when n is large, the distribution of $(\bar{Y} - \mu_Y)/\sigma_{\bar{Y}}$ is well approximated by the standard normal distribution. Here $\sigma^2_{\bar{Y}} = \frac{\sigma^2_Y}{n}$

Hence what we want to find out is

$$\Pr \left(\frac{\bar{Y} - \mu_Y}{\sqrt{\frac{\sigma^2_Y}{n}}} \leq \frac{\bar{y}^* - \mu_Y}{\sqrt{\frac{\sigma^2_Y}{n}}} \right)$$

where \bar{y}^* is our desired realization of the average \bar{Y} . This is Gaussian distributed!

$$\text{In the following, let } Z = \frac{\bar{Y} - \mu_Y}{\sqrt{\frac{\sigma^2_Y}{n}}}$$

We then solve the following way:

a) $\bar{y}^* = 0,43$ i.e. $\Pr(Z \geq \frac{0,43 - 0,4}{\sqrt{\frac{0,24}{100}}})$

i) $= 1 - \Pr(Z \leq \frac{0,43 - 0,4}{\sqrt{\frac{0,24}{100}}})$ (Note: we do this because standard cumulative Gaussian tables provides $\Pr(Z \leq z)$)

$$= 1 - \Pr(Z \leq 6,12)$$

$$= 1 - 0,7291 = \underline{\underline{0,2709}}$$

b) $\bar{y}^* = 0,37$ i.e. $\Pr(Z \leq \frac{0,37 - 0,4}{\sqrt{\frac{0,24}{400}}})$

$$\Pr(Z \leq -1,22) = \underline{\underline{0,1112}}$$

b) Want n to satisfy $\Pr(Z \leq \frac{0,41 - 0,4}{\sqrt{\frac{0,24}{n}}}) > -1,96$

and $\frac{0,39 - 0,4}{\sqrt{\frac{0,24}{n}}} < -1,96$ giving $n \geq 9220$

(calculation:
 $n > \frac{1,96}{0,01} \times 0,24 \approx 9220$)

2.24

$$Y_i \sim \text{iid } N(0, \sigma^2)$$

$$\text{a) } E\left(\frac{Y_i^2}{\sigma^2}\right) = \frac{1}{\sigma^2} E(Y_i^2) = \frac{\sigma^2}{\sigma^2} = 1$$

$$\text{Note that } \sigma^2 = \text{var}(Y_i) = E(Y_i^2) - \underbrace{[E(Y_i)]^2}_{=0} = E(Y_i^2)$$

b) We know that if

$Z \sim N(0, 1)$ then $Z^2 \sim \chi^2$ and generalized to sums of $N(0, 1)$ -vars.

$$\left(\frac{Y_i}{\sigma}\right) \sim N(0, 1)$$

Hence by definition

$$\sum_{i=1}^n \left(\frac{Y_i}{\sigma}\right)^2 = W \sim \chi_n^2$$

Since it's a sum of n independent $N(0, 1)$ vars.

$$\text{c) } E(W) = E\left(\frac{1}{\sigma^2} \sum_{i=1}^n Y_i^2\right) = \frac{1}{\sigma^2} E\left(\sum_{i=1}^n Y_i^2\right)$$

$$E(W) = \frac{1}{\sigma^2} \sum_{i=1}^n (E(Y_i^2)) = \frac{1}{\sigma^2} \sum_{i=1}^n \sigma^2 = \frac{n\sigma^2}{\sigma^2} = n$$

d)

By definition, let Z be standard normal and W chi-square with m degrees of freedom.

Let Z and W be independently distributed.

$$\text{Then } Z/\sqrt{W/m} \sim t_m$$

So what we need is

$$V = \frac{Y_1}{\sqrt{\frac{\sum_{i=2}^n Y_i^2}{n-1}}}$$

is for the numerator to be $N(0,1)$

and the denominator to be χ_{n-1}^2

We can easily do this by dividing both in the numerator and denominator by σ

$$V = \frac{(Y_1/\sigma)}{\sqrt{\frac{\sum_{i=2}^n (Y_i/\sigma)^2}{n-1}}}$$

where $(Y_1/\sigma) \sim N(0,1)$

$$\text{and } \sqrt{\frac{\sum_{i=2}^n (Y_i/\sigma)^2}{n-1}} \sim \chi_{n-1}^2$$

Hence by definition, $V \sim t_{n-1}$

2.26

Y_1, Y_2, \dots, Y_n have mean μ_Y , variance σ_Y^2
correlation ρ (for all pairs)

a)

We know that in general:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \text{var}(Y)}}$$

$$\Rightarrow \text{cov}(X, Y) = \text{corr}(X, Y) \sqrt{\text{var}(X) \text{var}(Y)}$$

$$\text{So: } \text{cov}(Y_i, Y_j) = \rho \sqrt{(\sigma_Y^2)^2} = \underline{\underline{\rho \sigma_Y^2}}$$

b) $E(\bar{Y}) = E\left(\frac{1}{n} \sum_{i=1}^n Y_i\right)$ where $n=2$

$$E(\bar{Y}) = E\left(\frac{1}{2} \sum_{i=1}^2 Y_i\right) = \frac{1}{2} \cdot (2\mu_Y) = \underline{\underline{\mu_Y}}$$

$$\textcircled{D} \quad \text{var}(\bar{Y}) = \text{var}\left(\frac{1}{2} \sum_{i=1}^2 Y_i\right)$$

$$\text{var}(\bar{Y}) = \frac{1}{4} \text{var}(Y_1 + Y_2)$$

$$\text{var}(\bar{Y}) = \frac{1}{4} [\text{var}(Y_1) + \text{var}(Y_2) + 2 \text{cov}(Y_1, Y_2)]$$

$$\textcircled{O} \quad \text{var}(\bar{Y}) = \frac{1}{4} [2\sigma_Y^2 + 2\rho\sigma_Y^2] = \underline{\underline{\frac{1}{2}\sigma_Y^2 + \frac{1}{2}\rho\sigma_Y^2}}$$

c) We can quite easily see that the formula

$$E(\bar{Y}) = E\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) \quad \text{quantifies to}$$

$$\textcircled{O} \quad \frac{1}{n} \sum_{i=1}^n (E(Y_i)) = \frac{1}{n} \sum_{i=1}^n \mu_Y = \frac{1}{n} n\mu_Y = \underline{\underline{\mu_Y}}$$

And recall now that since ρ is common to all pairs, $\text{cov}(Y_i, Y_j)$ is also common to all pairs so that

$$\begin{aligned} \text{var}\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) &= \frac{1}{n^2} [\underbrace{\text{var}(Y_1) + \text{var}(Y_2) + \dots + \text{var}(Y_n)}_{= n\sigma_Y^2} \\ &\quad + \underbrace{\text{cov}(Y_1, Y_2) + \text{cov}(Y_1, Y_3) + \dots + \text{cov}(Y_{n-1}, Y_n)}_{= (n-1)\rho\sigma_Y^2}] \end{aligned}$$

$$\text{Hence } \text{var}\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) = \text{var}(Y) = \frac{\sigma_Y^2}{n} + \frac{(n-1)}{n} \rho \sigma_Y^2$$

d) $\lim_{n \rightarrow \infty} \frac{\sigma_Y^2}{n} + \frac{(n-1)}{n} \rho \sigma_Y^2$

We see that the first term will vanish, while

$$\frac{n-1}{n} \quad \text{when } n \rightarrow \infty = 1$$

Hence $\lim_{n \rightarrow \infty} \text{var}(\bar{Y}) = \underline{\rho \sigma_Y^2}$

3. A

a) How to construct confidence intervals:

Let Z be a random variable such that

$$Z \sim \frac{\hat{X} - \mu}{\sigma/\sqrt{n}} \quad \text{then the probability that}$$

Z is in between two critical values can

$$\text{be written } \Pr(-z \leq Z \leq z) = 1-\alpha$$

where α is desired level of significance.

$$\text{Then } \Pr\left(-z \leq \frac{\hat{X} - \mu}{\sigma/\sqrt{n}} \leq z\right) = 1-\alpha$$

$$\Pr\left(-z \times \frac{\sigma}{\sqrt{n}} \leq \hat{X} - \mu \leq z \times \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$\boxed{\Pr\left(\hat{X} - z \times \frac{\sigma}{n} \leq \mu \leq \hat{X} + z \times \frac{\sigma}{n}\right) = 1-\alpha}$$

This is a general formula. In (most) of our applications, we use the slightly different notation:

$$\Pr\left(\hat{X} - t_{1-\alpha/2} \text{se}(\hat{X}) \leq \theta \leq \hat{X} + t_{1-\alpha/2} \text{se}(\hat{X})\right) = 1-\alpha$$

Then we can solve almost any problems with confidence intervals.

Difference in men's average hourly earnings between 1992 and 2008:

$$\bar{Y}_m^{2008} - \bar{Y}_m^{1992} = (24,98 - 23,27) = 1,71 = \theta_m$$

We then need $se(\bar{Y}_m^{2008} - \bar{Y}_m^{1992})$

$$se(\bar{Y}_m^{2008} - \bar{Y}_m^{1992}) = \sqrt{\frac{\sigma_{m,2008}^2}{n_{2008}} + \frac{\sigma_{m,1992}^2}{n_{1992}}}$$

$$= \sqrt{\frac{11,78^2}{1838} + \frac{10,17^2}{1594}} \approx 0,3747$$

Hence our 95% CI (i.e. $\alpha = 0,05$) is:

$$Pr(1,71 - 1,96 \times 0,3747 \leq \theta_m \leq 1,71 + 1,96 \times 0,3747) = 0,95$$

$$= [0,9756, 2,444]$$

Interpretation:

When constructing many intervals in this way, we believe that 95% of the time, we get an interval that contains the true value of θ_m . The probability of the true value lying inside a single confidence interval is either 0 or 1.

Hence we trust the theory (the method) and not the single confidence interval.

b) Same interval-type for women:

$$\bar{Y}_w^{2008} - \bar{Y}_w^{1992} = \theta_w = 20,87 - 20,05 = 0,82$$

$$se(\theta_w) = \sqrt{\frac{\sigma_{w,2008}^2}{n_{w,2008}} + \frac{\sigma_{w,1992}^2}{n_{w,1992}}}$$

$$se(\theta_w) = \sqrt{\frac{9,66^2}{1871} + \frac{7,87^2}{1368}} \approx 0,30846$$

$$\Rightarrow 95\% \text{ CI : } 0,82 \pm 1,96 \times 0,30846 = \underline{\underline{[0,215, 1,425]}}$$

c)

Since the observations are independent, we can use the same formulas as before.

$$\hat{\theta}_D = (\bar{Y}_{m, 2008} - \bar{Y}_{m, 1992}) - (\bar{Y}_{w, 2008} - \bar{Y}_{w, 1992}) = 1,71 - 0,82 = 0,89$$

$$se(\hat{\theta}_D) = \sqrt{\frac{11,78^2}{1838} + \frac{10,17^2}{1594} + \frac{9,66^2}{1871} + \frac{7,87^2}{1368}} \approx 0,48$$

Hence 95% CI: $0,89 \pm 1,96 \times 0,48 = \underline{[-0,0508, 1,8308]}$

3.20

$$S_{XY} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

$X_i, Y_i \sim \text{iid}$ First we can multiply and divide by n ($\frac{n}{n}$) to get:

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

Now follow procedure in appendix 3.3 and add and subtract μ_x and μ_y respectively to get:

$$\left(\frac{n}{n-1}\right) \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu_x + \mu_x - \bar{x})(Y_i - \mu_y + \mu_y - \bar{y}) \right)$$

Which rewrites into:

$$\left(\frac{n}{n-1}\right) \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu_x)(Y_i - \mu_y) \right) - \left(\frac{n}{n-1}\right) \left(\underbrace{\frac{1}{n} \sum_{i=1}^n (\bar{x} - \mu_x)(\bar{y} - \mu_y)}_{\text{only constants!}} \right)$$

$$= \left(\frac{n}{n-1}\right) \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu_x)(Y_i - \mu_y) \right) - \left(\frac{n}{n-1}\right) (\bar{x} - \mu_x)(\bar{y} - \mu_y)$$

When n grows large, $\bar{x} = \mu_x$ and $\bar{y} = \mu_y$ so the last term will vanish. (i.e. $\bar{x} \xrightarrow{P} \mu_x$, $\bar{y} \xrightarrow{P} \mu_y$)

Let now $W_i = (X_i - \mu_x)(Y_i - \mu_y)$ and note

that W_i is iid with mean σ_{xy} , and

second moment $E((X_i - \mu_x)^2(Y_i - \mu_y)^2)$

And by Cauchy-Schwarz Inequality

$$E((X_i - \mu_x)^2(Y_i - \mu_y)^2) \leq \sqrt{E(X_i - \mu_x)^4} \sqrt{E(Y_i - \mu_y)^4}$$

Hence the second moment of W_i is finite

Thus $\frac{1}{n} \sum_{i=1}^n w_i \rightarrow E(w_i) = \sigma_{xy}$

and it follows that $s_{xy} \xrightarrow{P} \sigma_{xy}$ since $\frac{n}{n-1} \rightarrow 1$

(3.21) We already made use of the fact that when \bar{Y}_m and \bar{Y}_w are independently distributed

$$se(\bar{Y}_m - \bar{Y}_w) = \sqrt{\frac{s_m^2}{n_m} + \frac{s_w^2}{n_w}} \quad (3.19)$$

$$\text{Now } n_m = n_w = n$$

So (3.19) will if we square it become

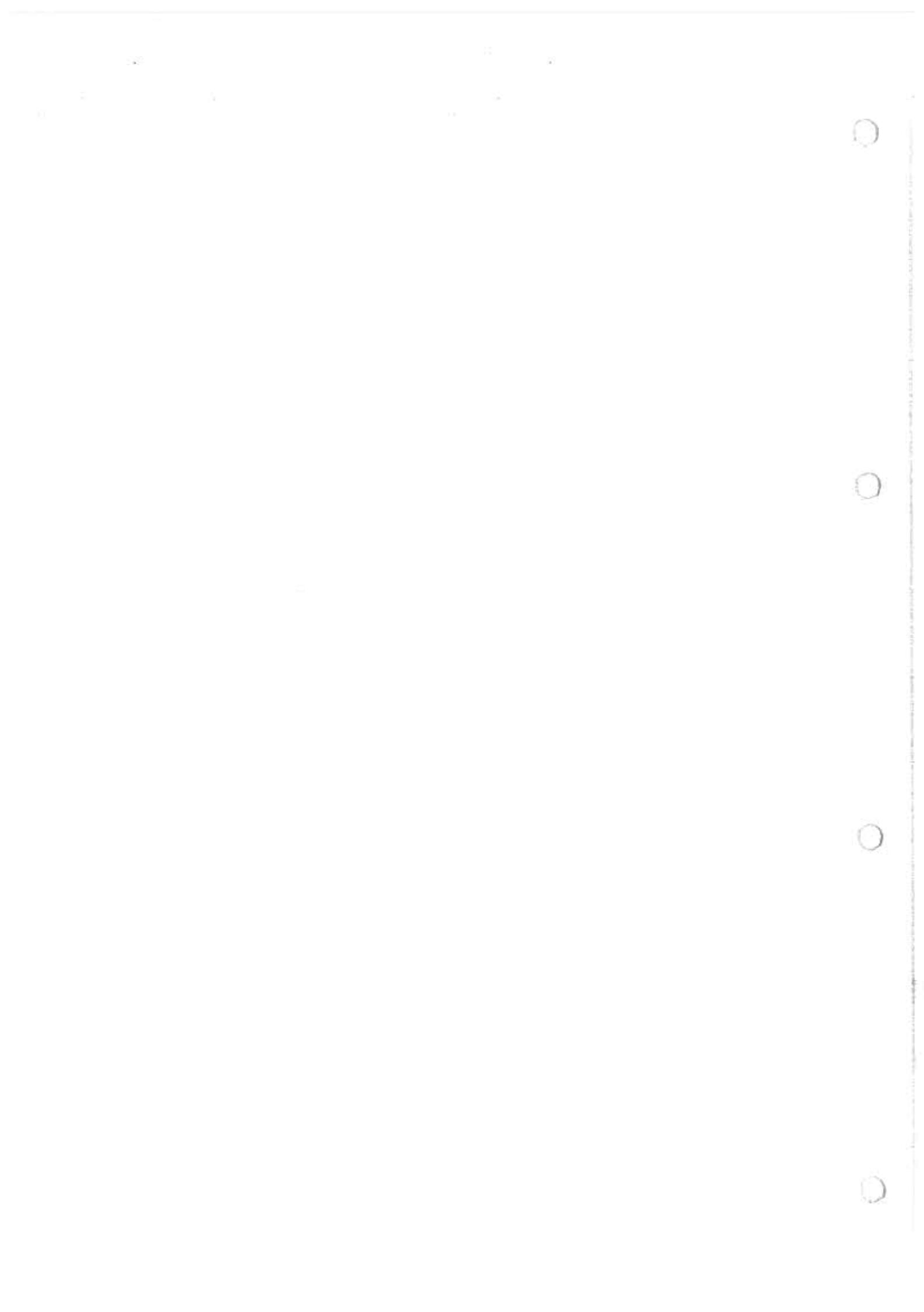
$$\begin{aligned} (se(\bar{Y}_m - \bar{Y}_w))^2 &= \frac{s_m^2}{n} + \frac{s_w^2}{n} \\ &= \frac{\frac{1}{n-1} \sum_{i=1}^n (Y_i^m - \bar{Y}_m)^2}{n} + \frac{\frac{1}{n-1} \sum_{i=1}^n (Y_i^w - \bar{Y}_w)^2}{n} \\ &= \frac{\sum_{i=1}^n (Y_i^m - \bar{Y}_m)^2 + \sum_{i=1}^n (Y_i^w - \bar{Y}_w)^2}{n(n-1)} \end{aligned}$$

Now using the pooled version:

$$[SE_{\text{pooled}}(\bar{Y}_m - \bar{Y}_w)]^2 = S_{\text{pooled}}^2 \times \frac{2}{n}$$

$$= \frac{1}{2(n-1)} \left[\sum_{i=1}^n (Y_i^m - \bar{Y}_m)^2 + \sum_{i=1}^n (Y_i^w - \bar{Y}_w)^2 \right] \frac{2}{n}$$

$$= \frac{\sum_{i=1}^n (Y_i^m - \bar{Y}_m)^2 + \sum_{i=1}^n (Y_i^w - \bar{Y}_w)^2}{n(n-1)}$$



SEMINAR 2

(4,5)

Y_i = points scored by i th student. ($0 \leq Y_i \leq 100$)

X_i = time available for i th student ($X_i = 90$ or 120)

Decided by a fair coin.

Model: $Y_i = \beta_0 + \beta_1 X_i + U_i$

a) U_i is the error term of the regression, and it measures "all other aspects" influencing points scored on the exam apart from available exam time. Of course different students have different abilities, skills, aptitude etc. which are factors that can affect exam scores.

b) $E(U_i | X_i) = 0$ is a key assumption when we do regression analysis. If it wasn't like this, then we would systematically have a biased estimator for the effect of available time on performance.

Because the assignment is random, u_i is independent of X_i (i.e. you don't get picked to have 120 min because you have a higher IQ than average etc).

Since $E(u_i) = 0$ due to u_i representing deviations from average, $E(u_i | X_i) = 0$ due to independency.

C) II) $(X_i, Y_i) \sim \text{iid}$ $i=1, 2, \dots, n$

III) Large outliers unlikely; X_i, Y_i have finite fourth moments.

II) is satisfied if this is a representative class, i.e. if you could pool all classes together and find the "average class" then it would resemble this class.

III) is satisfied since $Y_i \in (0, 100)$ and $X_i = 90$ or 120 so 4th order moments are finite.

d) Estimated regression: $\hat{Y}_i = 49 + 0,24X_i$

i) 90 min: $\hat{Y}_i = 49 + 0,24 \times 90 = \underline{\underline{70,6}}$

120 min: $\hat{Y}_i = 49 + 0,24 \times 120 = \underline{\underline{77,8}}$

○ 150 min: $\hat{Y}_i = 49 + 0,24 \times 150 = \underline{\underline{85}}$

ii) $\frac{dy}{dx} = 0,24 \Rightarrow 10 \text{ min extra gives } \underline{\underline{2,4}}$

4.6 $E(u_i | X_i) = 0$

○ We know that;

$$Y_i | X_i = \beta_0 + \beta_1 X_i | X_i + u_i | X_i \quad \text{by definition.}$$

and then $E(Y_i | X_i) = E(\beta_0 + \beta_1 X_i | X_i + u_i | X_i)$

$$= E(\beta_0) + E(\beta_1 X_i | X_i) + E(u_i | X_i) = \beta_0 + \beta_1 E(X_i | X_i)$$

$\Rightarrow E(Y_i | X_i) = \underline{\underline{\beta_0 + \beta_1 X_i}}$

(5.2)

$$\widehat{\text{Wage}} = 12,52 + 2,12 \times \text{male}$$

(0,23) (0,36)

$$R^2 = 0,06, \quad \text{SER} = 4,2$$

(standard error of regression).

$$\text{male} = \begin{cases} 0 & \text{if female} \\ 1 & \text{if male} \end{cases}$$

a) Gender gap: difference in mean earnings between men and women.

$$\text{DIM} = (12,52 + 2,12) - (12,52) = \underline{\underline{2,12}} \quad (\$/\text{hour})$$

b) We want to test if the coefficient in front of male (call it δ) is significantly different from 0.

○ Hence;

$$H_0: \delta = 0 \quad \text{vs.} \quad H_1: \delta \neq 0$$

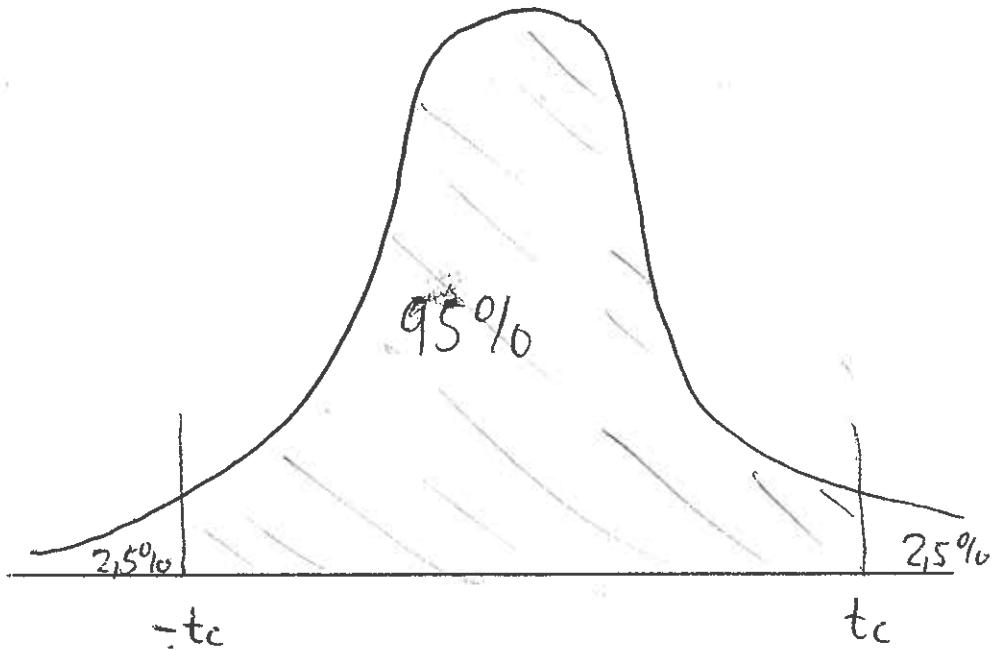
ABOUT THE TEST

Why is $H_0: \delta = 0$? Reason: because we always consider rejection-errors the worst!

○ "It is better to let some guilty convicts free, than imprison some unguilty persons".

Our secondary target is to make as few type II -errors as possible (i.e. not rejecting a false null-hypothesis). This is our desired level of

○ Significance. When we say we choose a "95% -significance level", it means we will only make rejection errors 5% of the time (type I errors). But choosing a higher significance level yields a trade-off since we then make type II -errors more often.



This figure shows that if we choose $\alpha = 0,05$ then 95% of the distribution will lead to not rejecting H_0 , and only if our observed value lies within the top (or bottom) 2,5%, will we reject H_0 . Note the symmetry of the t-distribution. Using a p-value to find out if one has an observation in the outer 5% is simply to check if $p\text{-value} < 0,05$ for which we reject H_0 .

Using the t-statistic to do the same, we use the fact that the t-distribution is symmetric and will have the criteria $|t_{\text{stat}}| > t_c$ where t_c is a critical value found in any t-table.

- How then to find t-stat and p-value?

t-stat is simply the measure of

"How many standard errors one is away from the desired point". In our case, how many standard errors away from 0 is 2,12 when the st. error is 0,36.

- Hence, in general:

$$t\text{stat} = \frac{\bar{t}_{\text{obs}} - \bar{t}}{se(\bar{t}_{\text{obs}})}$$

or in various forms
and notation.

In our particular case:

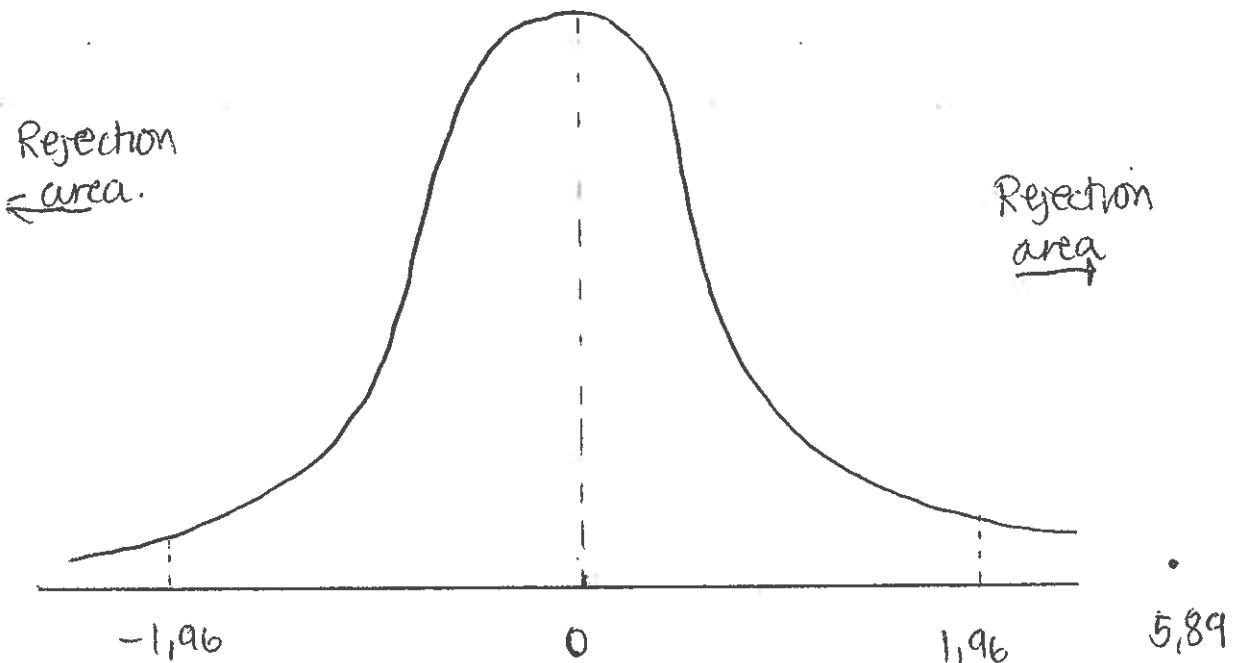
$$t\text{stat} = \frac{2,12 - 0}{0,36} \approx \underline{\underline{5,89}}$$

What is the critical value t_c ? Well since $n=250$

$$\text{we can use } t_{1-\frac{\alpha}{2}, n} = t_{0,975, \infty} = 1,96$$

$|t\text{stat}| = 5,89 > 1,96$ Hence we conclude that

we reject H_0 at 95% significance level, and we are unable to find evidence in data suggesting $\delta=0$



To find the p-value of this test, note that by the central limit theorem, the t-distribution converges in probability towards the standard normal when n is large, which means it follows the Gaussian (popular to denote Φ) cumulative distribution. Hence;

$$p\text{-value} = 2 \Phi(-|t_{\text{stat}}|) = 2 \Phi(-5,89)$$

$$p\text{-value} = 2 \times 0,0000 = \underline{\underline{0,000}}$$

The p-value gives the additional information that H_0 is rejected at any reasonable significance level.

- c) We went thoroughly through CI's last time. Using those results give:

$$95\% \text{ CI} = 2,12 \pm 1,96 \times 0,36 = \underline{\underline{[1,41, 2,83]}}$$

- d) If woman, mate = 0 so

$$\widehat{\text{wage}}_w = \underline{\underline{12,52 \text{ \$ / hour}}}$$

$$\widehat{\text{wage}}_m = 12,52 + 2,12 = \underline{\underline{14,64 \text{ \$ / hour}}}$$

- e) This should be a little intuitive:
 The only change is that the intercept will now be the male mean wage since female = 0 when the worker is male, and our δ from earlier must be the female-specific wage change.

Hence; $\widehat{\text{wage}} = 14,64 - 2,12 \times \text{female}$

What about R^2 and SER?

Consult the definitions!

Since the OLS residuals will be the same in the two regressions;

$$\text{wage} = \beta_0 + \beta_1 \times \text{male} + u_i$$

$$\text{wage} = \gamma_0 + \gamma_1 \times \text{female} + v_i$$

$$\Rightarrow \gamma_0 = \beta_0 + \beta_1 \quad \gamma_0 + \gamma_1 = \beta_0$$

so $\hat{u}_i = \hat{v}_i$, we get that

the sum of the squared residuals (SSR) is

$$\text{SSR} = \sum_i \hat{u}_i^2 \quad \text{and equal for both regressions.}$$

Which implies that;

$$\text{SER} = \left(\frac{\text{SSR}}{n-1} \right)^{\frac{1}{2}} \quad \text{and} \quad R^2 = 1 - \frac{\text{SSR}}{\text{SST}}$$

are equal in both regressions. Hence $R^2 = 0,06$

$$\text{SER} = 4,2$$

5.8

$$u_i \sim N(0, \sigma_u^2) \quad n = 30$$

$$\hat{Y} = 43,2 + 61,5X \quad R^2 = 0,54, \quad SER = 1,52$$

(10,2) (7,4)

a) 95% CI: $43,2 \pm t_{0,975,28} \times 10,2$

$$t_{0,975,30} \approx 2,05$$

↑
since $n=30$ and # of
estimated parameters = 2.

$$\Rightarrow 43,2 \pm 2,05 \times 10,2 = \underline{\underline{[22,29, 64,11]}}$$

b) $H_0: \beta_1 = 55$ vs. $H_1: \beta_1 \neq 55$ $\alpha = 0,05$

$$t_{\text{stat}} = \frac{61,5 - 55}{7,4} \approx 0,878$$

$$t_{0,975,28} = 2,05 \Rightarrow |t_{\text{stat}}| < t_c$$

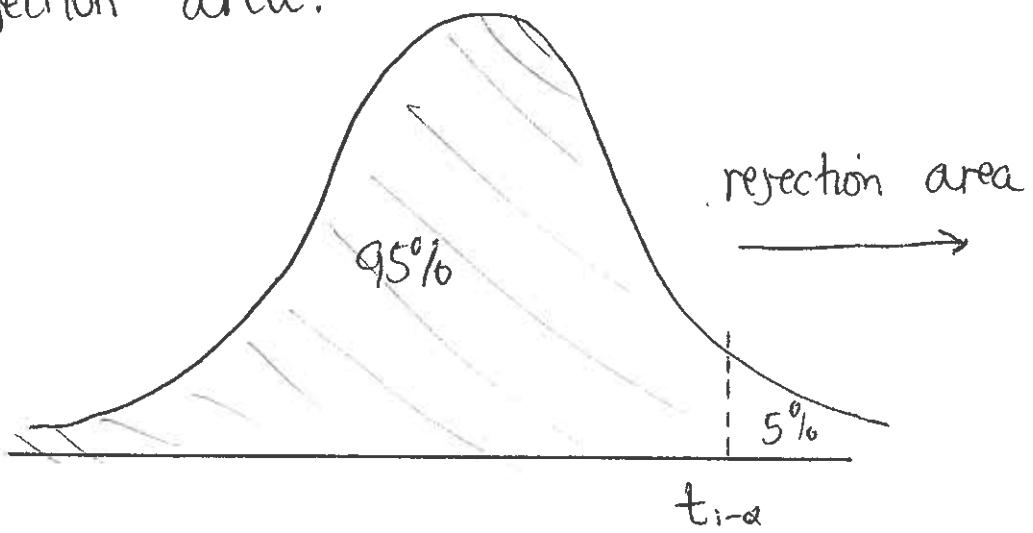
Hence do not reject H_0 at 95% significance level.

c)

$$H_0: \beta_1 = 55 \quad \text{vs.} \quad H_1: \beta_1 > 55$$

$$t_{\text{stat}} = 0,878$$

Only change is the critical value, since we now only consider the top 5% of the distribution as rejection area!



$$t_{0.95, 28} = 1,70.$$

Hence we do not reject H_0 at 95% significance level

Since $t_{\text{stat}} < t_c$.

Recall;

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$Y_i = \beta_0 + \beta_1 x_i + u_i$$

\bar{Y}_0 = sample mean when $x=0$

\bar{Y}_1 = sample mean when $x=1$

Let n_0 denote observations where $x=0$ and

n_1 observations where $x=1$. Then;

$$\sum_{i=1}^n x_i = n_1 \quad \text{and} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{n_1}{n}$$

$$\frac{1}{n_1} \sum_{i=1}^n x_i y_i = \bar{Y}_1$$

$$\text{Now; } \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i^2 - 2\bar{x}x_i + \bar{x}^2)$$

$$= \sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + n\bar{x}^2 = \sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + \bar{x} \sum_{i=1}^n x_i$$

$$= \sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i = n_1 - \frac{n_1^2}{n} = n_1 \left(1 - \frac{n_1}{n}\right)$$

$$= n_1 \left(\frac{n - n_1}{n} \right) = \frac{n_1 n_0}{n}$$

Then note that;

$$n_1 \bar{Y}_1 + n_0 \bar{Y}_0 = \sum_{i=1}^n Y_i$$

$$\Rightarrow \bar{Y} = \frac{n_1}{n} \bar{Y}_1 + \frac{n_0}{n} \bar{Y}_0$$

Now use the OLS formula for $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\text{and note that } \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum (x_i - \bar{x})y_i \\ = \sum (x_i y_i - \bar{y} n_1)$$

$$\Rightarrow \frac{\sum (x_i y_i - \bar{y} n_1)}{\frac{n_1 n_0}{n}} = \frac{\frac{n}{n_1} \sum (x_i y_i) - n \bar{y}}{n_0}$$

$$= \frac{n}{n_0} [\bar{Y}_1 - \bar{Y}] = \frac{n}{n_0} \left[\bar{Y}_1 - \frac{n_1}{n} \bar{Y}_1 - \frac{n_0}{n} \bar{Y}_0 \right]$$

$$= \frac{n}{n_0} \left[\frac{n-n_1}{n} \bar{Y}_1 - \frac{n_0}{n} \bar{Y}_0 \right] = \bar{Y}_1 - \bar{Y}_0$$

\therefore
 $= \frac{n_0}{n}$

Hence we proved that $\hat{\beta}_1 = \bar{Y}_1 - \bar{Y}_0$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \left(\frac{n_1}{n} \bar{Y}_1 + \frac{n_0}{n} \bar{Y}_0 \right) - (\bar{Y}_1 - \bar{Y}_0) \frac{n_1}{n}$$

$$\hat{\beta}_0 = \left(\frac{n_0}{n} + \frac{n_1}{n} \right) \bar{Y}_0 = \underline{\underline{\bar{Y}_0}}$$

E 4.1

a) Intercept = 1,082

Slope = 0,605

$$\Rightarrow \hat{a}_{he} = 1,082 + 0,605 \text{ age}$$

The slope tells us the unit increase, so

When workers age by 1 year, earnings increase by 0,605 \$/hour.

b) Bob: $1,082 + 0,605 \times 24 = \underline{\underline{16,81 \text{ \$/hour}}}$

Alexis: $1,082 + 0,605 \times 30 = \underline{\underline{19,23 \text{ \$/hour}}}$

c) Since R^2 is only 0,03 and R^2 is the measure of explained variation in the model, we conclude that age explains little of the variation in earnings across individuals.

E5.1

a) $H_0: \beta_1 = 0$ vs. $H_1: \beta_1 \neq 0$

Stata immediately gives us all we need!

$$t_{\text{stat}} = \frac{0,605 - 0}{0,0398} \approx \underline{\underline{15,18}}$$

$$|t_{\text{stat}}| > t_c \Rightarrow \text{reject } H_0.$$

5%: $t_c = 1,96 \Rightarrow \text{reject } H_0$

10%: $t_c = 1,645 \Rightarrow \text{reject } H_0$

1%: $t_c = 2,33 \Rightarrow \text{reject } H_0$

$$2 \Phi(-|t_{\text{stat}}|) = 2 \Phi(-15,18) = \underline{\underline{0,0000}}$$

Hence reject H_0 at any reasonable sign. level.

All of this of course provided immediately by Stata!

b) Provided immediately by Stata:

$$95\% \text{ CI} = 0,605 \pm 1,96 \times 0,0399 = \underline{\underline{[0,527, 0,683]}}$$

c) Use Stata output directly:

$$\text{p-value} = \underline{\underline{0,0000}} \Rightarrow \text{Reject}$$

$$d) \text{ p-value} = \underline{\underline{0,0006}} \Rightarrow \text{Reject.}$$

e) Difference in estimated $\hat{\beta}_1$ -coefficients

$$\text{is: } 0,92 - 0,30 = \underline{\underline{0,62}}$$

Standard error is:

$$se(\hat{\beta}_{1,\text{coll}} - \hat{\beta}_{1,\text{taugh}}) = \sqrt{(0,06^2 + 0,04^2)} = 0,07$$

$$\Rightarrow 95\% \text{ CI} = 0,62 \pm 1,96 \times 0,07 = \underline{\underline{[0,48, 0,76]}}$$

Hence yes, since the CI does not cover 0.

SEMINAR 3

6.5

n = 220 home sales in 2003

price = selling price in \$1000

BDR = number of bedrooms

Bath = " — bathrooms

Hsize = size of the house

Lsize = size of the lot in square feet

Age = age of the house

Poor = $\begin{cases} 1 & \text{if reported condition is "poor"} \\ 0 & \text{" — " "good"} \end{cases}$

$$\widehat{\text{Price}} = 119,2 + 0,485 \text{BDR} + 23,4 \text{Bath} + 0,156 \text{Hsize} + 0,002 \text{Lsize}$$

$$+ 0,090 \text{Age} - 48,8 \text{Poor}, \quad R^2 = 0,72 \quad SER = 41,5$$

a) $\frac{\partial \widehat{\text{PRICE}}}{\partial \text{Bath}} = 23,4 \Rightarrow \text{Increase in value} = \$23,400$

b) $\frac{\partial \widehat{\text{PRICE}}}{\partial \text{Bath}} + 100 \times \frac{\partial \widehat{\text{PRICE}}}{\partial \text{Hsize}} = 23,4 + 100 \times 0,156 = 23,4 + 15,6$

= 39 $\Rightarrow \text{Increase in value} = \$39,000$

c)

$$\frac{\partial \widehat{\text{PRICE}}}{\partial \text{PDR}} = -48,8 \Rightarrow -\$48,800$$

d) Adjusted R^2 (\bar{R}^2) is used to correct for the imprecise increase in R^2 when a new regressor is added. It is:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1} \frac{SSR}{SST}$$

whereas $R^2 = 1 - \frac{SSR}{SST}$

$$\Rightarrow \bar{R}^2 = 1 - \frac{n-1}{n-k-1} (1 - R^2)$$

$$\Leftrightarrow R^2 = 1 - \frac{n-k-1}{n-1} (1 - \bar{R}^2)$$

$$R^2 = 1 - \frac{220-6-1}{220-1} (1 - 0,72) \approx \underline{0,727}$$

6.10

(Y_i, X_{1i}, X_{2i}) satisfy Key Concept 6.4

i.e.: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$

1. $E(u_i | X_{1i}, X_{2i}) = 0$

2. $X_{1i}, X_{2i} \sim \text{iid}$

3. Finite 4th order moments

4. No perfect multicollinearity

And in addition: $\text{var}(u_i | X_{1i}, X_{2i}) = 4$

$$\text{var}(X_{1i}) = 6$$

$$n = 400$$

a)

Eq 6.17 in appendix 6.2:

$$\sigma_{\beta_1}^2 = \frac{1}{n} \left(\frac{1}{1 - \rho_{X_1, X_2}^2} \right) \frac{\sigma_u^2}{\sigma_{X_1}^2}$$

$$\Rightarrow \sigma_{\beta_1}^2 = \frac{1}{400} \times \frac{4}{6} = \underline{\frac{1}{600}} \quad (\approx 0.00167)$$

b) Assume $\rho_{x_1, x_2} = 0,5$

$$\Rightarrow \sigma_{\beta_1}^2 = \frac{1}{400} \left(\frac{1}{1 - 0,5^2} \right) \times \frac{4}{6} = \frac{1}{400} \times \frac{4}{3} \times \frac{4}{6} = \underline{\underline{\frac{1}{450}}} \\ (\approx 0,00222)$$

c) The statement correctly says that the variance of $\hat{\beta}_1$ would increase when the two regressors are correlated.

But leaving out x_2 from the regression could create an omitted variable bias, so we should not do that.

7.2

We went thoroughly through testing the last seminar, so we use our knowledge and take shortcuts now.

$$a) H_0: \beta_{\text{college}} = 0 \quad \text{vs. } H_1: \beta_{\text{college}} \neq 0$$

$$t_{\text{stat}} = \frac{5,46 - 0}{0,21} = 26,00$$

$$t_{0,997, 0,975} \approx 1,96 = t_c$$

$|t_{\text{stat}}| > t_c \Rightarrow \underline{\text{Reject } H_0}$ at 5% sign. level.

$$95\% \text{ CI: } 5,46 \pm 1,96 \times 0,21 = \underline{[5,05, 5,87]}$$

b) Let's solve this using CI - testing

$$H_0: \beta_{\text{female}} = 0 \quad \text{vs. } H_1: \beta_{\text{female}} \neq 0$$

$$\text{Construct CI: } -2,64 \pm 1,96 \times 6,20 = \underline{[-3,03, -2,25]}$$

Hence reject H_0 because the proposed value 0 does not lie within the 95% CI.

7.7

a) $H_0: \beta_{BDR} = 0$ vs. $H_1: \beta_{BDR} \neq 0$

$$t_{\text{stat}} = \frac{0,485 - 0}{2,61} \approx 0,186$$

$$t_{214, 0,975} \approx 1,96 = t_c$$

$|t_{\text{stat}}| < t_c \Rightarrow$ Do not reject H_0 at 5%

b) The conventional wisdom of houses with 5 bedrooms selling for much more than those with 2 holds since the coefficient β_{BDR} only captures the partial effects of an extra bedroom, holding all other variables constant. In reality, one would expect a 5-bedroom house to be much larger than a 2-bedroom house.

always

- c) Useful tip: start by constructing the 99% CI for the unit of the coefficient

$$t_{214, 0.995} \approx 2.58$$

$$\Rightarrow 0.002 \pm 2.58 \times 0.00048 = [0.0007616, 0.0032384]$$

o

\Rightarrow 2000 square feet 99% CI is

$$[2000 \times 0.0007616, 2000 \times 0.0032384] = [1,523.2, 6,476.8]$$

$$\Rightarrow [\underline{\underline{\$1523}}, \underline{\underline{\$6477}}]$$

o

- d) Probably using 1000 square feet as scale could make it easier to read the output since the estimated coefficient then would be 2 instead of 0.002.

o

e) Let's look deeper into F-testing than the problem proposed.

ABOUT F-TESTS

We might be interested in testing hypotheses with multiple conjectures, or so-called joint hypotheses.

Usually we use F-tests to do this (at least we do in ECON4150), and typically we want to test the significance of a model or parts of a model. In theory, we can have:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_R = 0 \quad \text{vs.} \quad H_1: \text{Any } \neq 0$$

H_1 can become a real mess, since any combination of the coefficients $\neq 0$ is a possibility. We can also test combinations of parameters, such as:

$$H_0: a\beta_1 + b\beta_2 = c\beta_3 \quad \text{vs.} \quad H_1: \text{not } H_0$$

And the tests can be one-sided:

$$H_0: \beta_2 + \beta_3 \leq 0 \quad \text{vs.} \quad H_1: \beta_2 + \beta_3 > 0$$

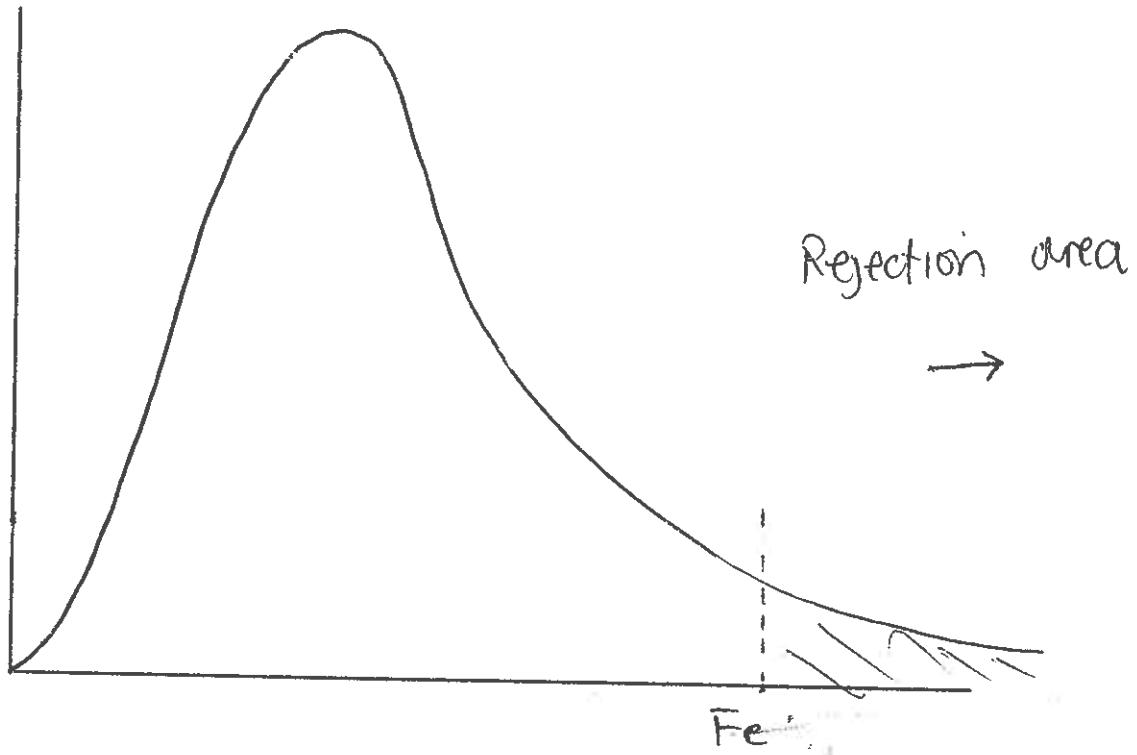
(Note that this latter becomes a t -test!)

- In general, we say that the model when H_0 is true is the restricted model, and when H_0 is false is the unrestricted model. We then compare sum of squared errors from these two models in the following (general) way:

$$F_{\text{stat}} = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$$

- Where J is the number of imposed restrictions (or rule of thumb: "number of equality signs in H_0 ")
- N is the total number of observations and K is the number of estimated parameters in the Unrestricted model. If the null is true, the F_{stat} has an F-distribution with J numerator degrees of freedom and $(N-K)$ denominator degrees of freedom. If the null is false, SSE_R and SSE_U differ a lot, which implies that the restrictions significantly reduce the ability of the model to

fit the data. "Too large" in this context is based on comparing with a critical value typically denoted $F_c = F_{1-\alpha, J, N-K}$ α being the desired significance level.



Note also that there exists an " R^2 -version" of the test that produces the exact same result:

$$F_{\text{stat}} = \frac{(R_u^2 - R_R^2) / J}{(1 - R_u^2) / (N - K)}$$

Now we know what we are answering in

7.7 e

$H_0: \beta_{BDR} = \beta_{Age} = 0$ vs. $H_1: \beta_{BDR} \neq 0$ or $\beta_{Age} \neq 0$
or both

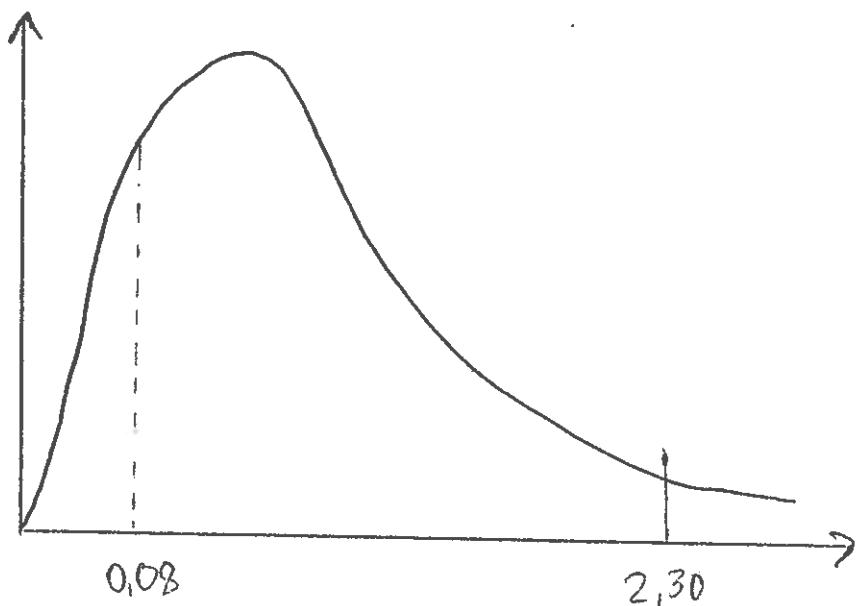
clearly $J=2$, $N=220$, $K=6$

$\alpha=0,1$.

They calculated F_{stat} for us. (too bad).

$F_c = 2,30$, $F_{\text{stat}} = 0,08$

Hence β_{BDR} and β_{Age} are not jointly significantly different from 0.



E6.2

a) Slope: -0,0733727

b) -0,0315387

c) Since the coefficient more than halved in value it is reasonable to believe it suffered from omitted variable bias.

d) Reg 1: $R^2 = 0,0074$, $\bar{R}^2 = 0,0072$

Reg 2: $R^2 = 0,2788$, $\bar{R}^2 = 0,2769$

Recall that $\bar{R}^2 = 1 - \frac{n-1}{n-k-1} \frac{SSR}{SST}$

and $R^2 = 1 - \frac{SSR}{SST}$

then realize that $n=3796 \Rightarrow \frac{n-1}{n-k-1} \rightarrow 1$

and hence $\bar{R}^2 \rightarrow R^2$

e)

$$\text{Dad Coll} = \begin{cases} 1 & \text{if dad has college education} \\ 0 & \text{otherwise} \end{cases}$$

The coefficient is positive and thus says there's a positive relationship between father's education and a person's education. If your father went to college, you on average complete 0.69 more years of education.

f) cue80: County unemployment rate in 1980

stwmfg80: State hourly wage in manufacturing in 1980

Hence these two capture the opportunity cost of going to college. When stwmfg80 increases, you forgo more wages, so college attendance should decline.

When cue80 increases, it becomes more difficult to find a job, so the opportunity cost decreases which should increase college attendance.

g)

$$\widehat{ed}_{BOB} = 8,828 - 0,032 \times 2 + 0,094 \times 58 + 0,145 \times 0 + 0,368 \times 1 \\ + 0,399 \times 0 + 0,395 \times 1 + 0,152 \times 1 + 0,696 \times 0 \\ + 0,023 \times 7,5 - 0,052 \times 9,75 = \underline{\underline{14,8}}$$

h) $\widehat{ed}_{JIM} = \underline{\underline{14,73}}$

E 7.3 a) The claim: $\beta_{DIST} = -0,15 / 2 = -0,075$

We have that the 95% CI is $[-0,100, -0,046]$

Hence since the claim is inside the 95% CI
it is consistent with the data.

b) Small changes in the coefficient when the number
of controls increase.

c) Yes. Both are positive and significant in regd.

SEMINAR 4

(8.2)

a) 500 sq. ft. addition to a house.

$$0,00042 \times 500 \times 100\% = \underline{\underline{21\%}}$$

95% CI (assume n "sufficiently large")

$$100\% \times 500 \times (0,00042 \pm 1,96 \times 0,000038) = \underline{\underline{[17,28\%, 24,72\%]}}$$

b) Since the dependant variable is identical in regression (1) and (2), we can use \bar{R}^2 as the measure of goodness of fit. Since

$$\bar{R}_{(1)}^2 = 0,72 < \bar{R}_{(2)}^2 = 0,74$$

$\ln(\text{size})$ is a better measure than size.

c) Pool = $\begin{cases} 1 & \text{if house has swimming pool} \\ 0 & \text{else} \end{cases}$

D \Rightarrow Effect on price of getting a pool (going from no pools to at least 1 pool) is $0,071 \times 100\% = \underline{\underline{7,1\%}}$

95% CI:

$$100\% \times [0,071 \pm 1,96 \times 0,034] = \underline{[0,436\%, 13,764\%]}$$

d)

An additional bedroom is estimated to increase house price by $100\% \times 0,0036 = \underline{\underline{0,36\%}}$

Statistically significant? $H_0: \beta_{BED} = 0$ vs. $H_1: \beta_{BED} \neq 0$

$$t_{\text{stat}} = \frac{0,0036 - 0}{0,037} = \underline{\underline{0,0973}}$$

$$t_c = 1,96 \Rightarrow |t_{\text{stat}}| < t_c \Rightarrow \text{Do not reject } H_0.$$

Evidence in data cannot support significance of β_{BED} .

The estimated effect should be small because this measures the effect of getting an extra bedroom without getting extra space, i.e. by converting other rooms into bedrooms.

e) $\ln(\text{size})^2$ appears only in regression (4)

and its test of significance is:

$$H_0: \beta_{\ln(\text{size})^2} = 0 \quad \text{vs.} \quad H_1: \beta_{\ln(\text{size})^2} \neq 0$$

$$t_{\text{stat}} = \frac{0,0078 - 0}{0,14} = 0,005571 < 1,96$$

Hence we cannot reject H_0 and consider $\ln(\text{size})^2$ unimportant.

f) Without view:

$$\frac{d \ln P_{\text{rice}}}{d \text{pool}} = \beta_{\text{pool}} + \beta_{\text{pool} \times \text{view}} \times \text{view}$$

$$\Rightarrow 0,071 + 0,0022 \times 0 = 0,071$$

$\Rightarrow \underline{7,1\%}$ increase in price

with view:

$$0,071 + 0,0022 \times 1 = 0,0732$$

$\Rightarrow \underline{7,32\%}$ increase in price

Is the difference significant?

$$H_0: \beta_{\text{pool} \times \text{view}} = 0 \quad \text{vs.} \quad H_1: \beta_{\text{pool} \times \text{view}} \neq 0$$

$$t_{\text{stat}} = \frac{0,0022 - 0}{0,1} = 0,022 < 1,96$$

\Rightarrow Do not reject H_0 . Not significant at 5% level.

⑧.4 a)

$$\Delta \ln \widehat{\text{AHE}} = \underbrace{\beta_x \bar{X} - \beta_x \bar{X}} + 0,0143 \times (3-2) - 0,000211 \times (3^2 - 2^2)$$

All unrelated to experience

$$\Delta \ln \widehat{\text{AHE}} = 0,0143 - 0,000211 \times 5 = \underline{\underline{0,013}} \quad (1,3\%)$$

$$b) \Delta \ln \widehat{\text{AHE}} = 0,0143 - 0,000211 \times \left(\frac{121}{11^2} - \frac{100}{10^2} \right)$$

$$\Delta \ln \widehat{\text{AHE}} = \underline{\underline{0,010}} \quad (1,0\%)$$

c) The results are different because the regression is non-linear in experience.

Hence changes depend on current level.

d) In (a) we expect the change to be

$$\beta_{\text{exp}} + 5\beta_{\text{exp}}^2$$

while in (b) we expect it to be

$$\beta_{\text{exp}} + 21\beta_{\text{exp}}^2$$

The difference between these is $16\beta_{\text{exp}}^2$

We thus need to check if β_{exp}^2 is significant at a 5% level.

$$H_0: \beta_{\text{exp}}^2 = 0 \quad \text{vs.} \quad H_1: \beta_{\text{exp}}^2 \neq 0$$

$$t_{\text{stat}} = \frac{-0,000211 - 0}{0,000023} \approx -9,17$$

$$t_c = t_{52782, 0,975} = 1,96$$

\Rightarrow Reject H_0 . The difference is supported by evidence in data.

e) No, this would only affect the level of lnAHE, but not the change associated with experience. These variables do not interact with experience

f) Then we could add the interaction-terms $p_{experience} \times female$ and $p_{experience^2} \times female$

$$\textcircled{8.7} \quad Female = \begin{cases} 1 & \text{if female} \\ 0 & \text{otherwise} \end{cases}$$

$$\widehat{\ln(\text{earnings})} = 6,48 - 0,44 \text{ female} \quad SER = 2,65$$
$$(0,01) \quad (0,05)$$

a) i) The estimated coefficient tells us that on average, females have $\ln(\text{earnings})$ 0,44 lower than men.

ii) SER tells us the standard deviation of the error term (in log points).

in its most general form

$$SER = \sqrt{s_{\hat{u}}^2} = \sqrt{\frac{\sum \hat{u}_i^2}{n-k-1}} = \sqrt{\frac{SSR}{n-k-1}}$$

- iii) Yes. Although many relevant characteristics (all of them actually) are left out that could hold explanatory power.
- iv) No, not necessarily. Discrimination would be the case if two workers are identical in all aspects except gender are paid different wages. To claim this, we need other control variables. This difference could simply be due to women being less educated, have physical burdens, less skilled, less productive etc.

b)

$$\widehat{\ln(\text{earnings})} = 3,86 - 0,28 \text{ female} + 0,37 \ln(\text{marketvalue}) \\ (0,03) \quad (0,04) \quad (0,004) \\ + 0,004 \text{ return} \\ (0,003)$$

$$n = 46,670 \quad \bar{R}^2 = 0,345$$

- i) If market value increases by 1%,
earnings increase by 0,37%.
- ii) Female is weakly correlated with at least one
of the added variables, and at least one of
them has significant explanatory power on
 $\ln(\text{earnings})$. Hence we had an omitted variable
bias in a).

- c) From the omitted variable bias formula:

$$\hat{\beta}_1 \rightarrow \beta_1 + \rho_{xu} \frac{\sigma_u}{\sigma_x} \quad \text{and assume the new } \hat{\beta}_1$$

is the true one. Then clearly since $\hat{\beta}_1 < \beta_1$
the bias must be negative, thus if we

① disregard "Return" whose coefficient is insignificant,
 female must be negatively correlated with
 in (marketvalue).

(8.10)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 (X_{1i} \times X_{2i}) + u_i$$

$$a) \Delta Y = f(X_1 + \Delta X_1, X_2) - f(X_1, X_2)$$

$$\Delta Y = \beta_1 \Delta X_1 + \beta_3 \Delta X_1 \times X_2$$

$$\Rightarrow \Delta Y = (\beta_1 + \beta_3 X_2) \Delta X_1$$

$$\frac{\Delta Y}{\Delta X_1} = \underline{\underline{\beta_1 + \beta_3 X_2}}$$

$$b) \Delta Y = f(X_1, X_2 + \Delta X_2) - f(X_1, X_2)$$

$$\Delta Y = \beta_2 \Delta X_2 + \beta_3 X_1 \Delta X_2$$

$$\Leftrightarrow \frac{\Delta Y}{\Delta X_2} = \underline{\underline{\beta_2 + \beta_3 X_1}}$$

c)

$$\Delta Y = f(x_1 + \Delta x_1, x_2 + \Delta x_2) - f(x_1, x_2)$$

$$\Delta Y = \beta_1 \Delta x_1 + \beta_2 \Delta x_2 + \beta_3 \Delta x_1 \Delta x_2 + \beta_3 \Delta x_1 x_2 + \beta_3 x_1 \Delta x_2$$

$$\underline{\Delta Y = (\beta_1 + \beta_3 x_2) \Delta x_1 + (\beta_2 + \beta_3 x_1) \Delta x_2 + \beta_3 \Delta x_1 \Delta x_2}$$

E8.1

a) The regression is linear, thus levels won't make a difference. If age increases by 1 year (no matter start age), earnings increase by

\$ 0,585

b) Function is still linear in age. Thus regardless of level, a one-year increase gives a 0,027 increase in ln age \Rightarrow 2,7% increase in earnings

c) Now we have non-linearity in age!

Age 25 to 26: ln age increase by $\ln 26 - \ln 25 \approx 0,0392$

Increase in ln age = $0,804 \times 0,0392 = 0,0315$

\Rightarrow earnings increase by 3,15%

Age 33 to 34: $\ln 34 - \ln 33 \approx 0,0299$

\Rightarrow Increase in ln age = $0,804 \times 0,0299 = 0,0240$

\Rightarrow earnings increase by 2,4%

d)

Age 25 to 26:

$$(0,081 \times 26 - 0,00091 \times 26^2) - (0,081 \times 25 - 0,00091 \times 25^2)$$

$$\approx 0,0346 \Rightarrow \text{earnings increase by } \underline{\underline{3,46\%}}$$

Age 33 to 34:

$$(0,081 \times 34 - 0,00091 \times 34^2) - (0,081 \times 33 - 0,00091 \times 33^2)$$

$$\approx 0,0200 \Rightarrow \text{earnings increase by } \underline{\underline{2,0\%}}$$

e) Since the two regressions share regressand and differ in the choice of one regressor, they can be compared on basis of their \bar{R}^2 . Regression c has a marginally higher \bar{R}^2 ($0,2005 > 0,2003$)

○ f)

In the d-regression we add age^2 to the regression. Since age^2 is not statistically significant ($p\text{-value} = 0,214$) we prefer the b-regression.

○

g) Again we can compare \bar{R}^2 .

\bar{R}^2 is marginally higher in regression C.

h) If we changed to females with BA

○ the plot would look similar, but would shift by the same amount as the coefficients on female and BA.

○

i)

$$\text{Alexis: } 0,081 \times 30 - 0,00091 \times 30^2 - 0,22 \times 1 + 0,40 \times 1 + 0,069 \approx \\ + 1,1 = \underline{\underline{2,96}}$$

$$\text{Jane: } \underline{\underline{2,49}} \quad \text{Bob: } \underline{\underline{3,11}} \quad \text{Jim: } \underline{\underline{2,71}}$$

Differences: Alexis - Jane = 0,469

$$\text{Bob - Jim} = \underline{\underline{0,40}}$$

j) We can actually do an F-test! (Yay!)

Restricted model: reg⁵

$$H_0: \beta_6 = \beta_7 = 0$$

$$H_1: \text{Not } H_0$$

Unrestricted model: reg⁶

$$F_{\text{stat}} = \frac{SSE_R - SSE_U | J}{SSE_U | N - K} \quad J = 2$$

$$F_{\text{stat}} = \frac{1695,49 - 1690,97 | 2}{1690,97 | 7703} \approx \underline{\underline{10,3}}$$

$$F_C = 4,61 \quad (\text{when } \alpha = 0,01)$$

\Rightarrow Reject H_0 .

k)

$$F_{\text{stat}} = \frac{1695,49 - 1690,74 / 2}{1690,74 / 7703} \approx \underline{10,82}$$

\Rightarrow Reject H₀.

l)

SEMINAR 5

Q. 9.2 $Y_i = \beta_0 + \beta_1 X_i + u_i$

$$\tilde{Y}_i = Y_i + w_i \quad , \quad w_i \sim \text{iID}$$

$$\tilde{Y}_i = \beta_0 + \beta_1 X_i + v_i$$

a)

By inserting for $Y_i = \beta_0 + \beta_1 X_i + u_i$ in \tilde{Y}_i ;

$$\tilde{Y}_i = \beta_0 + \beta_1 X_i + u_i + \underbrace{w_i}_{v_i}$$

$$\tilde{Y} = \beta_0 + \beta_1 X_i + v_i$$

$$\Rightarrow \underline{u_i + w_i} = v_i$$

b) Assume w_i independent of $Y_j, X_j \forall i, j$

and has finite 4th order moment.

Show: $E(v_i | X_i) = 0$, $X_i, \tilde{Y}_i \sim \text{iid}$,

X_i, \tilde{Y}_i have finite 4th order moments.

$$E(v_i | X_i) = E(u_i + w_i | X_i) = E(u_i | X_i) + E(w_i | X_i)$$

$$= 0 + 0 = \underline{0}$$

Since $\tilde{Y}_i = Y_i + w_i$ since $Y_i, w_i \sim i.i.d$

X_i, \tilde{Y}_j ($j \neq i$) independent since X_i is independent of both Y_j and w_j . Hence $(X_i, \tilde{Y}_i) \sim i.i.d$ from their joint distribution $\forall n$.

Since both (u_i, w_i) have finite 4th order moments and are mutually independent, v_i has finite 4th order moment too. Then it follows that X_i has finite 4th order moment.

c)

Whenever the least squares assumptions hold, OLS estimators are consistent.

d)

Again - whenever OLS assumptions hold everything is as normal. Inference may thus be conducted the regular way.

e)

If we assume that the measurement error is i.i.d. then error in the measure of X makes OLS assumptions fail. Thus we cannot trust inference and estimators are inconsistent.
If we have measurement error in Y , OLS is still valid, estimators are consistent and inference is valid. However, the variance of the regression will increase ($\sigma^2 = \sigma_u^2 + \sigma_w^2$)
Also if measurement errors are not i.i.d all of these results may change.

9.5

$$Q = \beta_0 + \beta_1 P + u \quad \text{DEMAND}$$

$$Q = \gamma_0 + \gamma_1 P + v \quad \text{SUPPLY}$$

$$E(u) = E(v) = 0, \quad \sigma_u^2, \quad \sigma_v^2, \quad \sigma_{uv} = 0$$

$$a) \quad \beta_0 + \beta_1 P + u = \gamma_0 + \gamma_1 P + v$$

$$\Rightarrow (\beta_1 - \gamma_1)P = \gamma_0 - \beta_0 + v - u$$

$$P = \frac{\gamma_0 - \beta_0 + v - u}{\beta_1 - \gamma_1}$$

$$\Rightarrow Q = \beta_0 + \beta_1 \frac{\gamma_0 - \beta_0 + v - u}{\beta_1 - \gamma_1} + u$$

$$Q = \frac{\beta_0 \beta_1 - \beta_0 \gamma_1 + \beta_1 \gamma_0 - \beta_1 \beta_0 + \beta_1 v - \beta_0 u + \beta_0 u - \gamma_1 u}{\beta_1 - \gamma_1}$$

$$Q = \frac{\beta_1 \gamma_0 - \beta_0 \gamma_1 + \beta_1 v - \gamma_1 u}{\beta_1 - \gamma_1}$$

$$b) E(P) = E \left[\frac{\gamma_0 - \beta_0 + v - u}{\beta_1 - \gamma_1} \right]$$

$$E(P) = \frac{1}{\beta_1 - \gamma_1} E(\gamma_0 - \beta_0 + v - u)$$

$$E(P) = \frac{1}{\beta_1 - \gamma_1} \left[E(\gamma_0) - E(\beta_0) + E(v) - E(u) \right] \\ = \gamma_0 - \beta_0 = 0 \quad = 0$$

$$\underline{E(P) = \frac{\gamma_0 - \beta_0}{\beta_1 - \gamma_1}}$$

$$E(Q) = E \left[\frac{\beta_1 \gamma_0 - \beta_0 \gamma_1 + \beta_1 v - \gamma_1 u}{\beta_1 - \gamma_1} \right]$$

$$E(Q) = \frac{1}{\beta_1 - \gamma_1} \left[\beta_1 \gamma_0 - \beta_0 \gamma_1 + E(\beta_1 v) - E(\gamma_1 u) \right] \\ = \beta_1 E(v) - \gamma_1 E(u) \\ = 0 \quad = 0$$

$$\underline{E(Q) = \frac{\beta_1 \gamma_0 - \beta_0 \gamma_1}{\beta_1 - \gamma_1}}$$

c)

$$\text{var}(P) = \text{var} \left(\frac{\gamma_0 - \beta_0 + v - u}{\beta_1 - \gamma_1} \right)$$

$$\text{var}(P) = \frac{1}{(\beta_1 - \gamma_1)^2} \text{var}(\gamma_0 - \beta_0 + v - u) \\ =_0 = 0$$

Note: $\text{var}(a) = 0$ when a constant.

$$\text{var}(ax + by) = a^2 \text{var}(x) + b^2 \text{var}(y) + 2ab \text{cov}(x, y)$$

$$\Rightarrow \text{var}(P) = \frac{1}{(\beta_1 - \gamma_1)^2} \left[\sigma_v^2 + \sigma_u^2 - 2 \text{cov}_{vu} \right] \\ = 0$$

$$\text{var}(P) = \frac{\sigma_v^2 + \sigma_u^2}{(\beta_1 - \gamma_1)^2}$$

$$\text{Var}(Q) = \frac{1}{(\beta_1 - \gamma_1)^2} \text{Var}[\beta_1 v - \gamma_1 u]$$

$$\text{Var}(Q) = \frac{1}{(\beta_1 - \gamma_1)^2} \left[\beta_1^2 \sigma_v^2 + \gamma_1^2 \sigma_u^2 - 2\beta_1 \gamma_1 \sigma_{uv} \right] = 0$$

$$\text{Var}(Q) = \frac{\beta_1^2 \sigma_v^2 + \gamma_1^2 \sigma_u^2}{(\beta_1 - \gamma_1)^2}$$

$$\text{Cov}(P, Q) = E[(P - E(P))(Q - E(Q))]$$

$$= E \left[\left(\frac{\gamma_0 - \beta_0 + v - u}{\beta_1 - \gamma_1} \right) \left(\frac{\beta_1 v_0 - \beta_0 \gamma_1 + \beta_1 v - \gamma_1 u}{\beta_1 - \gamma_1} \right) \right]$$

$$= 0 \quad = 0$$

$$= E \left[\frac{\beta_1 v^2 - \beta_1 v u - \gamma_1 v u + \gamma_1 u^2}{(\beta_1 - \gamma_1)^2} \right]$$

$$= \frac{\beta_1 \sigma_v^2 + \gamma_1 \sigma_u^2}{(\beta_1 - \gamma_1)^2}$$

d)

i) We know that by OLS

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} = \frac{\text{cov}(X, Y)}{\text{var}(X)} \quad \text{in general.}$$

$$\Rightarrow \hat{\beta}_1 = \frac{\text{cov}(P, Q)}{\text{var}(P)} = \frac{\beta_1 \sigma_v^2 + \gamma_1 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = E(Y) - \hat{\beta}_1 E(X) \quad \text{in general}$$

$$\Rightarrow \hat{\beta}_0 = E(Q) - \hat{\beta}_1 E(P)$$

$$\hat{\beta}_0 = \frac{\beta_1 \gamma_0 - \beta_0 \gamma_1}{\beta_1 - \gamma_1} - \frac{\beta_1 \sigma_v^2 + \gamma_1 \sigma_u^2}{\sigma_v^2 + \sigma_u^2} \frac{\gamma_0 - \beta_0}{\beta_1 - \gamma_1}$$

$$\begin{aligned} \hat{\beta}_0 &= \frac{\beta_1 \gamma_0 (\sigma_v^2 + \sigma_u^2) - \beta_0 \gamma_1 (\sigma_v^2 + \sigma_u^2) - \beta_1 \gamma_0 \sigma_v^2 + \beta_0 \gamma_1 \sigma_u^2}{(\beta_1 - \gamma_1) (\sigma_v^2 + \sigma_u^2)} \\ &\quad + \frac{\beta_0 \beta_1 \sigma_v^2 - \gamma_1 \gamma_0 \sigma_u^2}{(\beta_1 - \gamma_1) (\sigma_v^2 + \sigma_u^2)} \end{aligned}$$

$$= \frac{\beta_1 \gamma_0 \sigma_u^2 - \beta_0 \gamma_1 \sigma_v^2 + \beta_0 \beta_1 \sigma_v^2 - \gamma_1 \gamma_0 \sigma_u^2}{(\beta_1 - \gamma_1) (\sigma_v^2 + \sigma_u^2)}$$

$$= \frac{\gamma_0 (\beta_1 - \gamma_1) \sigma_u^2 + \beta_0 (\beta_1 - \gamma_1) \sigma_v^2}{(\beta_1 - \gamma_1) (\sigma_v^2 + \sigma_u^2)}$$

$$\hat{\beta}_0 = \frac{\gamma_0 \sigma_u^2 + \beta_0 \sigma_v^2}{\sigma_v^2 + \sigma_u^2}$$

ii) Find $\hat{\beta}_1 - \beta_1$ where β_1 is true value.

$$\hat{\beta}_1 - \beta_1 \xrightarrow{P} \frac{\beta_1 \sigma_v^2 + \gamma_1 \sigma_u^2}{\sigma_v^2 + \sigma_u^2} - \beta_1$$

$$\hat{\beta}_1 - \beta_1 \xrightarrow{P} \frac{\beta_1 \sigma_v^2 + \gamma_1 \sigma_u^2 - \beta_1 \sigma_v^2 - \beta_1 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$$

$$\hat{\beta}_1 - \beta_1 \xrightarrow{P} \frac{(\gamma_1 - \beta_1) \sigma_u^2}{\sigma_v^2 + \sigma_u^2} > 0 \quad , \text{ i.e. too large.}$$

Since γ_1 is associated with supply (hence it's positive) and β_1 with demand (hence it's negative).

9.10

Internal validity:

Statistical inferences about causal effects are valid for the population being studied.

This means first of all that the estimator of this causal effect should be unbiased and consistent. Secondly, inference in general should have the desired significance level.

External validity

Inferences and conclusions can be generalized from the population studied to other populations and settings. This means that differences in populations may be a threat to external validity, so can differences in settings.

The Return to Education and the Gender Gap

Internal validity concerns:

- in the last paragraph, potential omitted variables is mentioned. This could cause biased estimates and false inference.
- Since the experiment was done using full-time workers only, we can have a selection bias.

External validity concerns:

- Relative demand^{and supply} for skilled and unskilled workers in the economy may change over time. This can affect the returns to education.

E.8.3

a) Dist increases from $2 \rightarrow 3$, $6 \rightarrow 7$ (linear!)
 ED reduced by $\underline{-0,0366}$ years.

b) In ed drops by $-0,00260$ for both (linear!)
 $\Rightarrow \Delta ed = \underline{-0,26\%}$

c) From $2 \rightarrow 3$:

$$(-0,081 \times 3 + 0,0046 \times 3^2) - (-0,081 \times 2 + 0,0046 \times 2^2) = \underline{-0,058}$$

From $6 \rightarrow 7$:

$$(-0,081 \times 7 + 0,0046 \times 7^2) - (-0,081 \times 6 + 0,0046 \times 6^2) = \underline{-0,0212}$$

d) (3) is preferred to (1) since the coefficient on
 dist² is statistically significant.

$$t_{\text{stat}} \approx \frac{0,0046 - 0}{0,0023} \approx \underline{2}$$

e) i)

The only change for a white male would be a change in the intercept.

ii) It seems the quadratic equation becomes positive for $\text{dist} > 10$

There are 44 observations with $\text{dist} > 10$ (1% of the sample). Hence it is quite imprecise.

f) Coefficient: -0,366. This is the effect when both parents completed college.

h) Regression on pg. 6 adds interaction terms.

$H_0: \beta_6 = \beta_7 = 0$ vs. $H_1: \text{Not } H_0$

$$F_{\text{stat}} = \frac{\frac{8924,08}{-8912,80 / 2}}{8912,80 / 3781} = 2,39$$

Hence they are significant at a 10% level with $F_c = 2,36$, but not at a 5% level with $F_c = 3,00$

E9.3

a) INTERNAL VALIDITY

- Omitted variables: Students from wealthier families might live closer to colleges and have higher average years of completed education.
- Misspecification of functional form: we saw in E8.3. that the choice is ambiguous.
- Errors-in-variables: Such as income and ownhome which are imperfect measures of wealth.
- Sample selection: Unlikely to be a problem here.
- Simultaneous causality: Parents who want kids to go to college locate closer to the college.
- Inconsistency of standard errors: Not likely to be a problem.

b) EXTERNAL VALIDITY

Note first the difference in observations
(the former having 3796, latter 943)

This means standard errors in the West-set will be approximately twice those of non-West.

$$\text{ratio} \approx \sqrt{\frac{n_{\text{non-west}}}{n_{\text{west}}}} \approx 2$$

Q

Q

Q

Q

SEMINAR 6

10.2

a) $y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_1 D_{1i} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + u_{it}$

$$D_{1i} = 1 \text{ if } i=1, 0 \text{ otherwise}$$

$$D_{2i} = 1 \text{ if } i=2, 0 \text{ otherwise}$$

$$D_{3i} = 1 \text{ if } i=3, 0 \text{ otherwise}$$

Hence since for any $i \in \{1, 3\}$ only one binary regressor is equal to 1 we have that

$$D_{1i} + D_{2i} + D_{3i} = 1 = X_{0,it} \quad \forall i \in \{1, 3\}$$

b) Again, for each $i \in \{1, n\}$ only one binary regressor is equal to 1, so

$$D_{1i} + D_{2i} + \dots + D_{ni} = 1 = X_{0,it} \quad \forall i \in \{1, n\}$$

c) Including all binary variables and the constant causes perfect multicollinearity. OLS estimators cannot be computed in this case.

10.5

$$Y_{it} = \beta_1 X_{1,it} + \alpha_i + \lambda_t + u_{it} \quad i = 1, 2, \dots, n \\ t = 1, 2, \dots, T$$

$$\Leftrightarrow Y_{it} = \beta_0 + \beta_1 X_{1,it} + \delta_2 B_{2t} + \dots + \delta_T B_{Tt} \\ + \gamma_2 D_{2i} + \dots + \gamma_n D_{ni} + u_{it}$$

$$\text{Again: } D_{2i} = 1 \text{ if } i = 2$$

$$D_{3i} = 1 \text{ if } i = 3 \quad \text{etc. . .}$$

$$B_{2t} = 1 \text{ if } t = 2$$

$$B_{3t} = 1 \text{ if } t = 3 \quad \text{etc. . .}$$

$$\text{Then } \beta_0 = \alpha_1 + \lambda_1 \quad \text{and} \quad \gamma_i = \alpha_i = \alpha_1$$

$$\delta_t = \lambda_t - \lambda_1$$

Since when for instance $i = t = 2$ then the first model is $Y_{22} = \beta_1 X_{1,22} + \alpha_2 + \lambda_2 + u_{22}$

$$\text{second is: } Y_{22} = \beta_0 + \beta_1 X_{1,22} + \delta_2 + \gamma_2 + u_{22}$$

$$\Rightarrow Y_{22} = \beta_0 + \beta_1 X_{1,22} + \lambda_2 - \lambda_1 + \alpha_2 - \alpha_1 + u_{22}$$

$$Y_{22} = \beta_1 X_{1,22} + \lambda_2 + \alpha_2 + \underbrace{\beta_0 - \lambda_1 - \alpha_1}_{= -\beta_0} + u_{22}$$

$$Y_{22} = \beta_1 X_{1,22} + \lambda_2 + \alpha_2$$

10.7

- a) Note that average snow fall will not vary over time! She collected a lot of data, then estimated the average, then added the average as a variable. It will thus be perfectly collinear with the state fixed effect.
- b) Now since the researcher has a variable that varies over time, this can be used alongside the state fixed effects.

10.11

$$\text{Eq (10.22)} \quad \hat{\beta}_i^{\text{DM}} = \frac{\sum_{i=1}^n \sum_{t=1}^T (X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i)}{\sum_{i=1}^n \sum_{t=1}^T (X_{it} - \bar{X}_i)^2}, \quad \bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}$$

$$\hat{\beta}_i^{\text{BA}} = \frac{\sum_{i=1}^n (X_{i2} - X_{i1})(Y_{i2} - Y_{i1})}{\sum_{i=1}^n (X_{i2} - X_{i1})^2}$$

Show: $T=2$ gives $\hat{\beta}_i^{\text{BA}} = \hat{\beta}_i^{\text{DM}}$

$$\begin{aligned}
 \text{Then } \hat{\beta}_1^{\text{DM}} &= \frac{\sum_{i=1}^n \sum_{t=1}^2 (x_{it} - \frac{1}{2} \sum_{t=1}^2 x_{it})(y_{it} - \frac{1}{2} \sum_{t=1}^2 y_{it})}{\sum_{i=1}^n \sum_{t=1}^2 (x_{it} - \frac{1}{2} \sum_{t=1}^2 x_{it})^2} \\
 \hat{\beta}_1^{\text{DM}} &= \frac{\sum_{i=1}^n \left(\frac{1}{4} (x_{i2} - x_{i1})(y_{i2} - y_{i1}) + \frac{1}{4} (x_{i2} - x_{i1})(y_{i2} - y_{i1}) \right)}{\sum_{i=1}^n \left(\frac{1}{4} (x_{i2} - x_{i1})^2 + \frac{1}{4} (x_{i2} - x_{i1})^2 \right)} \\
 \hat{\beta}_1^{\text{DM}} &= \frac{\sum_{i=1}^n (x_{i2} - x_{i1})(y_{i2} - y_{i1})}{\sum_{i=1}^n (x_{i2} - x_{i1})^2} = \hat{\beta}_1^{\text{BA}}
 \end{aligned}$$

$$\begin{aligned}
 * \text{ Since: } & \sum_{i=1}^n \sum_{t=1}^2 (x_{it} - \frac{1}{2} \sum_{t=1}^2 x_{it})(y_{it} - \frac{1}{2} \sum_{t=1}^2 y_{it}) \\
 &= \sum_{i=1}^n \left[(x_{i1} - \frac{1}{2}(x_{i1} + x_{i2})) \left(y_{i1} - \frac{1}{2}(y_{i1} + y_{i2}) \right) \right. \\
 &\quad \left. + (x_{i2} - \frac{1}{2}(x_{i1} + x_{i2})) \left(y_{i2} - \frac{1}{2}(y_{i1} + y_{i2}) \right) \right] \\
 &= \left[\left(-\frac{1}{2}(x_{i2} - x_{i1}) \right) \left(-\frac{1}{2}(y_{i2} - y_{i1}) \right) + \left(\frac{1}{2}(x_{i2} - x_{i1}) \right) \right. \\
 &\quad \left. \left(\frac{1}{2}(y_{i2} - y_{i1}) \right) \right] = \left[\frac{1}{4} (x_{i2} - x_{i1})(y_{i2} - y_{i1}) + \right. \\
 &\quad \left. \frac{1}{4} (x_{i2} - x_{i1})(y_{i2} - y_{i1}) \right]
 \end{aligned}$$

b)

Evidently, the coefficient on shall drops a lot.

It is now -0,046.

This suggests important omitted variable bias in reg2! The new coefficient is not statistically significantly different from 0.

c) The coefficient drops even further and is not statistically significantly different from 0.

It is now -0,028. As it seems, the Fstat for time effects is 21,62 (see do-file!)

Hence we prefer the model with time effects.

d) Again it seems as if the initially seemingly significant results are spurious. All the qualitative results are the same.

E10.1

a)

i) The coefficient is $-0,368$ which suggests that shall-issue laws reduce violent crimes by 36%. This is a quite large real-world effect.

ii) In reg1 the coefficient was $-0,443$ and its confidence interval includes the coefficient in reg2. Both are statistically significant.

iii) • Attitudes towards guns and crime
• Quality of the police-force
• Crime-prevention programs.

e)

Threats to internal validity:

- Potential two-way causality between this year's incarceration rate and the number of crimes.
- Two-way causality between crimes and shall.

- page 4.

- f) From the fourth regression (reg 4) we can see that the 95% CI for shall is (-11%, 53%) which covers 0. We can also note that the p-value is 0,495 which is very high. Hence there is no statistical evidence that concealed weapons laws have any effect on crime rates.

SEMINAR 8

12.2

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

- a) $Z_i = X_i$ means that a regression of X on Z gives the coefficient 1.
Hence the instrument enters the first stage regression. By Key Concept 4.3 we know that $\text{corr}(X_i, u_i) = 0$ which implies that $\text{corr}(Z_i, u_i) = 0$ thus Z_i exogenous.
- b) There are no w's, so (1) is satisfied
Since Key Concept 4.3 holds, we know that (X_i, Z_i, Y_i) are iid from their joint distribution since $X_i = Z_i$, thus (2) holds.
Since X_i, Y_i have finite 4th order moment, so does Z_i , thus (3) holds.
(4) was shown in (a).

c)

$$\hat{\beta}_1^{\text{TSLS}} = \frac{s_{zy}}{s_{zx}} \quad \text{by definition (eq 12.4)}$$

Since $Z = X$ we immediately have that

$$\hat{\beta}_1^{\text{TSLS}} = \frac{s_{zy}}{s_{zx}} = \frac{s_{xy}}{s_x^2} = \hat{\beta}_1^{\text{OLS}}$$

12.6

We use the homoskedastic-only F-statistic
(which is only valid if $u_i \sim \text{homoskedastic}$)

where we simply check if there is any
statistical causality from X to Z .

$$F_{\text{stat}} = \frac{(0,05 - 0) / 1}{0,95 / 98} \approx \underline{5,16}$$

$F_C \approx 6,63$ at a 1% level for which
the instrument is weak.

With $n = 400$ we get

$$F_{\text{stat}} = \frac{(0,05 - 0) / 1}{0,95 / 498} \approx 26,2$$

Hence the instrument is not weak at any reasonable significance level.

12.8

a) $Q_i^s = \beta_0 + \beta_1 P_i + u_i^s$

$$Q_i^d = \gamma_0 + u_i^d$$

u_i^s, u_i^d mutually independent and iid.

$$E(u_i^s) = E(u_i^d) = 0$$

$$Q_i^s = Q_i^d \Rightarrow \beta_0 + \beta_1 P_i + u_i^s = \gamma_0 + u_i^d$$

$$\Rightarrow P_i = \frac{\gamma_0 - \beta_0 + u_i^d - u_i^s}{\beta_1}$$

Now the formula for covariance is

$$\text{cov}(P_i, u_i^s) = E(P_i u_i^s) - E(P_i)E(u_i^s)$$

$$\text{And by assumption } E(u_i^s) = 0 \Rightarrow \text{cov}(P_i, u_i^s) = E(P_i u_i^s)$$

$$E(P_i u_i^s) = E\left[\frac{\gamma_0 u_i^s - \beta_0 u_i^s + u_i^d u_i^s - (u_i^s)^2}{\beta_1}\right]$$

Since u_i^d, u_i^s are mutually independent;

$$\begin{aligned} \text{cov}(u_i^s, u_i^d) &= E(u_i^d u_i^s) - E(u_i^d)E(u_i^s) = 0 \\ &= 0 &= 0 &= 0 \end{aligned}$$

$$\begin{aligned} \Rightarrow E(P_i u_i^s) &= \frac{\gamma_0}{\beta_1} E(u_i^s) - \frac{\beta_0}{\beta_1} E(u_i^s) + \frac{1}{\beta_1} E(u_i^s u_i^d) \\ &\quad - \frac{1}{\beta_1} E[(u_i^s)^2] = 0 \end{aligned}$$

$$\text{Now; } \text{var}(u_i^s) = \sigma_{u_i^s}^2 = E[(u_i^s)^2] - [E(u_i^s)]^2 = E[(u_i^s)^2]$$

$$\text{Hence; } \text{cov}(P_i, u_i^s) = -\frac{\sigma_{u_i^s}^2}{\beta_1} \neq 0$$

=====

b) Let's derive the result:

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n} \sum_{i=1}^n (P_i - \bar{P}) u_i^s}{\frac{1}{n} \sum_{i=1}^n (P_i - \bar{P})^2} \quad * \quad (\text{Derived from } \hat{\beta}_1^{\text{OLS}})$$

$$\Rightarrow \hat{\beta}_1 \xrightarrow{P} \beta_1 + \frac{P \rho u^s \sigma_{u^s} \sigma_P}{\sigma_P^2}$$

As we can see, as long as P and u^s are correlated $\hat{\beta}_1$ is inconsistent since it does not converge in probability towards β_1 .

* Derivation: $\hat{\beta}_1 = \frac{\sum (P_i - \bar{P})(Q_i - \bar{Q})}{\sum (P_i - \bar{P})^2}$

$$\hat{\beta}_1 = \frac{\sum (P_i - \bar{P})(\beta_1(P_i - \bar{P}) + (u_i^s - \bar{u}^s))}{\sum (P_i - \bar{P})^2}$$

$$\hat{\beta}_1 = \frac{\beta_1 \sum (P_i - \bar{P})^2 + \sum (P_i - \bar{P}) u_i^s - \bar{u}^s \sum (P_i - \bar{P})}{\sum (P_i - \bar{P})^2} \stackrel{?}{=} 0$$

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n} \sum (P_i - \bar{P}) u_i^s}{\frac{1}{n} \sum (P_i - \bar{P})^2}$$

(c) We would need an instrumental variable that is uncorrelated with P , but uncorrelated with U .

We can calculate γ_0 directly using OLS and then use Q as an instrument since Q is uncorrelated with shifts in supply. We find γ_0 by regressing Q on only a constant term (regressor = 1) and we then find average quantity.

12.10

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + U_i$$

No data available on W_i .

a) $\text{cov}(Z_i, W_i) = 0$

$$\hat{\beta}_1^{\text{IV}} = \frac{\text{cov}(Z_i, Y_i)}{\text{cov}(Z_i, X_i)} = \frac{\text{cov}(Z_i, \beta_0 + \beta_1 X_i + \beta_2 W_i + U_i)}{\text{cov}(Z_i, X_i)}$$

$$\frac{\beta_1 \text{cov}(z_i, x_i) + \beta_2 \text{cov}(z_i, w_i)}{\text{cov}(z_i, x_i)}$$

$$= \beta_1 + \frac{\beta_2 \text{cov}(z_i, w_i)}{\text{cov}(z_i, x_i)}$$

Hence only if $\text{cov}(z_i, w_i) = 0$ will $\hat{\beta}_1$ be consistent.

b) From a) the answer is "inconsistent"

E12.2 a) Women with more than two kids will on average work 5,39 weeks less

b) Since both fertility and weeks worked are choice variables, morekids is likely to be correlated with the error term.

For instance; a woman who works more on average may also be a woman who is less likely to have another child in the first place. $\beta_{morekids}$ positively biased.

c) Second regression shows that if the two first children are of same gender, the couple is 6,75% more likely to have a third child. The effect is statistically significant ($t\text{-value} = 35,19$, $p\text{-value} = 0,0000$).

d) Samesex is unrelated to any other variables in the model, including the error term. Thus it is exogenous. We can also see from the second regression that the F-statistic is 1238 so the instrument is relevant.

e) No, same argument as in d.

f) The effect is -6,314, suggesting that women with more than 2 children work 6,314 weeks less on average.

g) No, there are no important changes. This is because all the other regressors are unrelated to Samesex, so they were not omitted variables.

SEMINAR 9

B.3

- a) Average treatment effect: The mean of the individual causal effects in the population under study. Mathematically we denote it:

$$ATE = E(Y_i | X_i=1) - E(Y_i | X_i=0)$$

or the difference between the expected outcome of the treatment group and the control group respectively.

$$\Rightarrow ATE = 1241 - 1201 = \underline{40}$$

- b) We would have non-random assignment if, for either men or women, the probability of being selected to the control group or the treatment group differed. Hence for random assignment we need $\hat{p}_{men} = \hat{p}_{women} = 0,5$. We can test

first for men: $H_0: \hat{p}_{men} = 0,5$ vs. $H_1: \hat{p}_{men} \neq 0,5$

$$t_{stat} = \frac{\hat{p}_{men} - 0,5}{\sqrt{\frac{1}{n_{men}} \hat{p}_{men}(1 - \hat{p}_{men})}} = \frac{0,55 - 0,5}{\sqrt{\frac{1}{100} 0,55(1 - 0,55)}} \approx 1,00$$

Since $n = 100$ we have $|t_{cl}| \approx 2$

Hence we cannot reject H_0 of random assignment.

For women we get similarly;

$$t_{stat} = \frac{0,45 - 0,5}{\sqrt{\frac{1}{100} 0,45(1-0,45)}} \approx -1,00$$

Hence we conclude the same for women.

13.5 A little bit generally on threats to internal validity:

- Failure to randomize, typically based on characteristics or preferences
- Failure to follow treatment protocol, typically when people misunderstand the experiment or fail to comply.
- Attrition, or dropping out of the study
- Experimental effects, or behavioral changes
- Small samples

a) This is a clear example of attrition, which is a threat to internal validity. The remaining subjects are representative only for the population that excludes male athletes. Note that attrition only affects the internal validity of the experiment if male athletes differ in their average causal effect from the population excluding them.

b) Clearly this group of engineering students fail to comply completely to the protocol. This is a threat to internal validity. This will bias the OLS estimator of the average causal effect.

c) The art students' choice of not learning how to use their accounts (or failure to learn it) is no threat to internal validity. The experiment was focused on the causal effect of the availability of internet on grades, not the use of it.

d) This is similar to b, partial compliance.

(13.7) $T = 2 \quad (t = 1, 2)$

$X_{it} = 1$ for $t = 2$ for individuals in treatment

$X_{it} = 0$ otherwise

(*) $Y_{it} = \alpha_i + \beta_1 X_{it} + \beta_2 (D_t \times W_i) + \beta_0 D_t + V_{it}$

α_i fixed effects (individual)

D_t binary variable $= \begin{cases} 1 & \text{if } t=2 \\ 0 & \text{if } t=1 \end{cases}$

$$\Delta Y_i = Y_{i2} - Y_{i1}$$

Derive,

$$(13.6) \quad \Delta Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_{1i} + \dots + \beta_{1+r} W_{ri} + u_i$$
$$i = 1, 2, \dots, n$$

So we immediately from (*) get:

$$Y_{i2} - Y_{i1} = \Delta Y_i = \beta_1 (X_{i2} - X_{i1}) + \beta_2 (D_2 - D_1) W_i + \beta_0 (D_2 - D_1) + (V_{2i} - V_{1i})$$

Define $X_{i2} - X_{i1} \equiv X_i$, $V_{2i} - V_{1i} \equiv U_i$

$$\underline{\Delta Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i}$$

Which is (13.6) in the case of a single W regressor.

(B.8) $Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 G_i + \beta_3 D_t + U_{it}$

$$G_i = \begin{cases} 1 & \text{if individual } i \text{ is in treatment group} \\ 0 & \text{if individual } i \text{ is in control group} \end{cases}$$

Eq. (13.4) $\hat{\beta}_1^{\text{did}} = (\bar{Y}_{\text{treat-after}} - \bar{Y}_{\text{treat-before}}) - (\bar{Y}_{\text{cont-after}} - \bar{Y}_{\text{cont-before}})$

$$= \Delta \bar{Y}_{\text{treat}} - \Delta \bar{Y}_{\text{control}}$$

We get that if an individual is;

a) in treatment group, after treatment:

then $t = 2$ so $D_t = X_{it} = 1$ and $G_i = 1$

$$\Rightarrow \bar{Y}_{it}^{\text{treat-after}} = \beta_0 + \beta_1 + \beta_2 + \beta_3$$

b) in treatment group, before treatment:

$t = 1$ so $D_t = X_{it} = 0$ and $G_i = 1$

$$\Rightarrow \bar{Y}_{it}^{\text{treat-before}} = \beta_0 + \beta_2$$

c) in control group, after treatment

then $t = 2$ so $D_t = 1$ and $X_{it} = 0$, $G_i = 0$

$$\Rightarrow \bar{Y}^{\text{control-after}} = \beta_0 + \beta_3$$

d) in control group, before treatment:

then $t = 1$ so $D_t = 0$, $X_{it} = 0$, $G_i = 0$

$$\Rightarrow \bar{Y}^{\text{control-before}} = \beta_0$$

$$\Rightarrow \hat{\beta}_1^{\text{DID}} = [(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2)] - [(\beta_0 + \beta_3) - \beta_0]$$

$$\underline{\hat{\beta}_1^{\text{DID}}} = \beta_1$$

E13. 1

a) Call-back rates for whites implies black = 0

$$\Rightarrow \underline{0,0965} \quad \text{or} \quad \approx 9,7\%$$

For African Americans implies black = 1

$$\Rightarrow 0,0965 - 0,032 = \underline{0,0645} \approx 6,5\%$$

The 95% CI for this difference equals the 95% CI for the coefficient on black which is reported to be $[-0,0472949, -0,0167708]$

It does not cover 0, so the effect is statistically significant ($t = -4,11$)

b) Adding the interaction term female * black shows that we cannot conclude that there are gender-differences connected to being white or African American ($p\text{-value} = 0,544$) hence we do not reject the null of no statistical significance.

- c) Using the third regression we see that call-back rates are 0,01406 higher for high quality resumes. ($p\text{-value} = 0,071$) For white applicants we use regression 4 and see that it is 0,023. For African American applicants it is $0,023 - 0,018 = \underline{\underline{0,005}}$. The difference between African Americans and whites is thus $0,005 - 0,023 = \underline{\underline{-0,018}}$ with $p\text{-value} = \underline{\underline{0,253}}$
- d) In the table at page 3 and 4 we can see that only the difference in computer skills and call-back is significant.

ABOUT PROBABILITY LIMITS

In order to say what happens to a stochastic quantity as $n \rightarrow \infty$ we use probability limit (or plim). In general, if $a(y^n)$ is some function of y^n and the plim of $a(y^n)$ as $n \rightarrow \infty$ is a^* we write

$$\underset{n \rightarrow \infty}{\text{plim}} a(y^n) = a^* \quad \text{Or more formally;}$$

$$\lim_{n \rightarrow \infty} \Pr(|a(y^n) - a^*| < \epsilon) = 1$$

which means that the probability that $a(y^n)$ differs from a^* by less than an infinitesimal ϵ will converge to 1 as $n \rightarrow \infty$.

Example: Suppose y_t is coin tosses taking the value 1 if heads and 0 if tails.

After n tosses the proportion of heads is

$$p(y^n) = \frac{1}{n} \sum_{t=1}^n y_t \quad \text{given a fair, unbiased coin.}$$

it comes as no surprise that $E(y_t) = 1/2$ and

$$\underset{n \rightarrow \infty}{\text{plim}} p(y^n) = \frac{1}{2}$$

Example

Suppose \bar{x} is the sample mean of x_t
 $t = 1, 2, \dots, n$, a sequence of random variables
with expectation μ . Provided x_t is independent

$$\text{plim}_{n \rightarrow \infty} \bar{x} = \text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n x_t = \mu$$

SEMINAR 10

14.2

a) $y_t = 1200 \times \ln(1P_t / 1P_{t-1})$

When the monthly change is small, the statement is correct. The monthly change in

$1P$ is $\frac{1P_t - 1P_{t-1}}{1P_{t-1}} \times 100$ which can be

$$\text{approximated by } [\ln 1P_t - \ln 1P_{t-1}] \times 100 = 100 \ln \left(\frac{1P_t}{1P_{t-1}} \right)$$

We convert this into annual change by multiplying with 12: $1200 \times \ln \left(\frac{1P_t}{1P_{t-1}} \right)$

b) $\hat{y}_t = 1,377 + 0,318 y_{t-1} + 0,123 y_{t-2} + 0,068 y_{t-3} + 0,001 y_{t-4}$

2000: 8	2000: 9	2000: 10	2000: 11	2000: 12
y_t : 8,55	2,60	-2,52	-3,67	-7,36

$$y_{2000:12} = 1200 \times \ln \left(\frac{147,300}{148,200} \right) \approx -7,36$$

$$\hat{y}_{2001:1} = 1,377 + 0,318 \times (-7,36) + \dots + 0,001 \times 8,55$$

$$\hat{y}_{2001:1} = \underline{\underline{-1,58}}$$

c)

The test for this is:

$$H_0: \gamma_{t-12} = 0 \quad \text{vs.} \quad H_1: \gamma_{t-12} \neq 0 \quad \text{choose } \alpha = 0,05$$

$$t_{\text{stat}} = \frac{\hat{\gamma}_{t-12} - 0}{\text{se}(\hat{\gamma}_{t-12})} = \frac{-0,054}{0,053} \approx -1,0189$$

$41 \times 12 = 492$ no. of obs.

And since T is large we have $t_c = 1,96$

$$\Rightarrow |t_{\text{stat}}| < t_c \Rightarrow \underline{\text{Do not reject } H_0}$$

e) Again, $T = 492$

$$\text{Formulas: } BIC(p) = \ln \left(\frac{SSR(p)}{T} \right) + (p+1) \frac{\ln T}{T}$$

$$AIC(p) = \ln \left(\frac{SSR(p)}{T} \right) + (p+1) \frac{2}{T}$$

About BIC:

The SSR will necessarily decrease as number of added lags grow. The second term however will increase when lags are added. The number of lags that minimize this BIC is a consistent estimator for the true lag length.

About AIC:

The first term is exactly the same. Only difference in the second term is that $\ln T$ is switched for 2. This is usually smaller than $\ln T$, meaning a smaller decrease in SSR is needed in AIC to justify more lags.

	1	2	3	4	5	6
$\frac{\text{SSR}(p)}{T}$	59,30	58,00	57,71	57,71	57,68	57,55
$\ln \left(\frac{\text{SSR}(p)}{T} \right)$	4,0826	4,0605	4,0554	4,0553	4,0549	4,0527
$(p+1) \frac{\ln T}{T}$	0,0252	0,0378	0,0504	0,0630	0,0756	0,0882
$(p+1) \frac{2}{T}$	0,0081	0,0122	0,0163	0,0203	0,0244	0,0285

	1	2	3	4	5	6
BIC	4,1078	4,0983	4,1058	4,1183	4,1305	4,1409
AIC	4,0907	4,0727	4,0717	4,0757	4,0793	4,0812

Hence BIC is smallest for $T=2$ and AIC is smallest for $T=3$. Q

(14.3)

$$\widehat{\Delta \ln(1P_t)} = 0,061 + 0,00004t - 0,018 \ln(1P_{t-1}) \\ + 0,333 \Delta \ln(1P_{t-1}) + 0,162 \Delta \ln(1P_{t-2})$$

About the Augmented Dickey-Fuller test Q

We want to test $H_0: \delta = 0$ vs. $H_1: \delta < 0$

$$\text{in } \Delta Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + u_t$$

Under the null, Y_t has a stochastic trend,
while under H_1 Y_t is stationary. Q

We use the normal t-statistic to test this!

But, we can? add a deterministic time trend (a "t")

$$\Rightarrow Y_t = \beta_0 + \alpha t + \gamma Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + u_t$$

Note that the ADF statistic does not have a normal distribution even in large samples.

$$ADF_{stat} = \frac{-0,018}{0,007} = -2,57$$

We have intercept and time trend so we see that at a 10% level the critical value is -3,12 and because $-2,57 > -3,12$ we do not reject H_0 . Note that in ADF we reject at low values.

b)

The results do support the specification in 14.2
Why? Since the series in 14.3 are non-stationary
so-called "random walk". Consider a variable
 y_t behaving as a random walk:
 $y_t = y_{t-1} + v_t$, non-stationary with stochastic
trend. Then we can render the model stationary
by taking the first difference:

$$\Delta y_t = y_t - y_{t-1} = v_t$$

The variable is said to be first-difference stationary
(or integrated of order 1 $\rightarrow I(1)$)

14.4

Granger causality: If lagged values of a random variable has a causal effect on another random variable, we say it is Granger causing it.

$$Y_{t-1} \rightarrow X_t, \text{ i.e. in } X_t = \pi_{21} Y_{t-1} + \pi_{22} X_{t-1} + \epsilon_t$$

$\pi_{21} \neq 0$ we say that Y is Granger-causing X .

The Granger causality F-statistic tests the hypothesis that the coefficients on all the lags of one variable is zero. In;

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + \delta_1 X_{t-1} + \dots + \delta_k X_{t-k}$$

We test; $H_0: \delta_1 = \dots = \delta_k = 0$ vs $H_1: \text{Not } H_0$.

(We could add several different X 's!)

a) The F-statistic is 2,35 and we have

$$H_0: \delta_1 = \delta_2 = \delta_3 = \delta_4 = 0 \quad \text{vs. } H_1: \text{Not } H_0$$

so $J = 4$ (numerator degrees of freedom)

$$N - K + 1 = 492 - 9 + 1 = 484 \Rightarrow F_c = 2,37$$

Hence we cannot reject H_0 at 5% significance level.

This means we cannot reject the null that the interest rates have no predictive content for IP growth.

b) F-statistic is 2,87

And ; $J = 4$, $N - K + 1 = 479$

$$\Rightarrow F_C = 2,37 \Rightarrow \text{Reject } H_0.$$

E 14.1 a) $\bar{\Delta Y_t} = 0,0077219$
 $(0,0092819)$

b) $\overline{\Delta Y_{t, \text{annual, percent}}} = 3,088771$
 $(3,712768)$

d) Unit free by construction!

E 14.2

a) Coefficient is 0,343

p-value = 0,000

95% CI: [0,198 , 0,489]

b) p-value 0,186 \rightarrow not significant.

Only slightly higher \bar{R}^2

We prefer AR(1).

c) Minimized BIC when no. of lags is : 1
Minimized AIC when no. of lags is : 2

E 14.5

a) $\bar{R}^2_{ADL} = 0,173 \quad \bar{R}^2_{AR} = 0,115$

b) p-value = 0,0006

