**UNIVERSITY OF TORONTO**
**Faculty of Arts and Science**

**APRIL 2017 EXAMINATIONS**

**ECO220Y1Y**

**Duration - 3 hours**

**Examination Aids: A non-programmable calculator**

**U of T E-MAIL:** ______________________________@MAIL.UTORONTO.CA

| | |
|---|---|
| **SURNAME (LAST NAME):** | |
| **GIVEN NAME (FIRST NAME):** | |
| **UTORID: (e.g. LIHAO118)** | |

This exam includes these ***10 pages*** plus the *Supplement*. There are ***8 questions*** (some with multiple parts) with varying point values worth a total of ***120 points***. You are responsible for turning in all 10 pages of this exam and for writing your name and identifying information above *before* the end of the exam.

The *Supplement* is 12 pages and contains graphs, tables, and other information needed to answer some of the exam questions as well as the aid sheets (formula sheets and Normal, $t$ and $F$ statistical tables). ***Anything written on the Supplement will NOT be graded. Once the exam begins, carefully detach the Supplement.***

Write your answers clearly, completely and concisely in the designated space provided immediately after each question. An answer guide for your response ends each question: it is underlined so you do not miss it. It lets you know what is expected: for example, a quantitative analysis (which shows your work and reasoning), a fully-labelled graph, and/or sentences. Anything requested by the question and/or guide is required. If the answer guide does not request sentences, provide only what is requested (e.g. quantitative analysis). For questions with multiple parts (e.g. (a) – (c)), attempt each part even if you have trouble with some. Be careful with parts marked [No partial credit]: your answer and work must be completely correct.

***Your entire answer must fit in the designated space provided immediately after each question.*** No extra space/pages are possible. You *cannot* use blank space for other questions nor can you write answers on the *Supplement*. ***Write in PENCIL and use an ERASER as needed.*** This way you can make sure to fit your final answer (including work and reasoning) in the appropriate space. Most questions give more blank space than is needed to answer. ***Follow the answer guides and avoid excessively long answers.***

**(1)** [10 pts] At some universities, all incoming students write an entrance test. Results can be reported as scores (e.g. total points earned) or as percentiles. For example, a student could be told that s/he scored 162 out of 250 possible points or that s/he is at the 55.38th percentile. Regardless of whether scores are skewed, bi-modal, or Normal, the percentile results follow the Uniform distribution, taking values between 0 and 100. Suppose all incoming students are randomly assigned to dormitories. For a dormitory with 38 students, what is the probability that *the average percentile result* is above 64? <u>Answer with a quantitative analysis and illustrate the probability with a fully-labelled graph of the sampling distribution of the sample mean.</u>

**(2)** See the ***Supplement for Question (2)****: Big Mac Prices in China.*

**(a)** [6 pts] *Fully interpret* ***both*** the OLS intercept and slope coefficients in the Excel regression output. Write interpretations that would be clear to someone who has *not* read the *Supplement*. Answer with 2 precise sentences that specify the units of measurement and clearly explain what those two numbers mean in a practical sense.

**(b)** [6 pts] If the variable t were measured in years, not months, since June 2005, what would the equation for the OLS line be? For example, if instead of a value of 19 months, t took a value of 1.5833 years. Also, would the $R^2$ be higher, lower or the same? Explain your answers. Answer with an OLS equation using standard notation and 2 – 3 sentences.

**(c)** [6 pts] Find the number 4.30887E-17 in the Excel output labelled "ANOVA" under "*Significance F*." Is that number extremely huge or extremely tiny? Why should someone just looking at a scatter diagram (of local price versus t) expect such an extreme number? Which conclusion does this extreme number support? What is the relevant null hypothesis stated *in words*, not formal notation? <u>Answer with 3 – 4 sentences.</u>

**(d)** [8 pts] Imagine an isolated data-entry mistake: the local price in June 2005 is recorded as 19.50 instead of 10.50. Would this data-entry mistake create an outlier? What effect, if any, would this mistake have on the number 0.448771056 under "*Regression Statistics*" in the Excel output? Explain *why* 0.448771056 would go up, down or remain the same, making sure to include its units of measurement and what that number measures in a practical sense. <u>Answer with 4 – 5 sentences.</u>

**(3)** See the ***Supplement for Question (3)****: Learning by Doing in Automobile Manufacturing.*

**(a)** [5 pts] *Fully interpret* the number -0.289, which appears in Table 1, Column (1), Panel A. Write an interpretation that would be clear to someone who has *not* read the *Supplement*. Answer with 1 precise sentence that clearly explains what that number means in a practical sense.

**(b)** [3 pts] How many regressions are reported in Table 1? Which of these correspond to Figure 2? Answer with a number of regressions and the column number and panel that goes with Figure 2.

**(c)** [3 pts] By the last time period of the available data, approximately how many cars had the manufacturer produced in total during the study period? Answer with a number *and* your work. [No partial credit]

**(d)** [5 pts] Compute the ***99%*** confidence interval estimate of the learning rate using the results in Table 1, ***Column (2), Panel A****.* Answer with a quantitative analysis.

**(e)** [6 pts] Table 1 reports the usual test of statistical significance for each coefficient at a 5% significance level (in other words, the statistical test that standard software packages, like Excel, automatically conduct). All are marked with a "*" as being statistically significant. For the "*" next to 0.007 in ***Column (2), Panel A***, what are: the formal hypotheses being tested, the value of the test statistic, and the relevant critical value(s)? <u>Answer with formal hypotheses in standard notation and the other requested values.</u>

**(f)** [8 pts] Consider ***Column (2), Panel A*** of Table 1. What does the positive coefficient on the time trend mean? Would it still be positive if the cumulative production variable were dropped? (In other words, if this empirical model $\ln(D_t) = \alpha + \gamma * t + \varepsilon_t$ were estimated instead.) Explain. Justify your answer by referencing relevant evidence from the *Supplement for Question (3)*. <u>Answer with 3 – 5 sentences.</u>

**(4)** See the ***Supplement for Question (4)****: Pilots and Fuel Efficient Flights.*

**(a)** [14 pts] *After* the experiment, *is there a difference* between the control group and the combined treatment groups (i.e. all three treatment groups combined)? Include a quantitative measure of the strength of the evidence. Is the difference statistically significant? If so, at which significance levels? Is the *point estimate* of the difference economically significant (in other words, is it far enough from zero that the airline and policy makers should be interested)? <u>Answer with formal hypotheses in standard notation, a quantitative analysis, the P-value, and 2 – 4 sentences.</u>

**(b)** [3 pts] Levitt and List (2011) discuss the "Hawthorne Effect" where people change their behavior simply because they know they are being studied by researchers. There is strong evidence of the Hawthorne Effect in this study: the captains in the control group had their efficiency improve by about the same amount as those in the treatment groups when we compare the period before the experiment with the period after the experiment. This is true even though the researchers didn't do any intervention for the pilots in the control group (other than tell them they were being monitored). Hence, it makes sense to look at the last column of Table 4 that combines all four groups. <u>Correctly fill in the FOUR blanks below. [No partial credit]</u>

According to Table 4, the percent of ALL flights (i.e. flights by any of the captains) that are efficient is ______ percent before the experiment versus ______ percent after the experiment. This is a ______ percentage point increase and a ______ percent increase.

**(5)** [4 pts] Suppose air pollution – measured by the Air Pollution Index (API) – across days in a city is Normally distributed with a mean of 75 and a standard deviation of 20. For a randomly selected day, what is the probability that the API is between 100 and 120? <u>Answer with a number *and* your work. [No partial credit]</u>

**(6)** [4 pts] A charity is engaged in a fundraising campaign. Each day volunteers contact a fresh set of potential donors. The amount of money raised each day is *independent* of the other days. Suppose the expected amount of money raised per day is $20,000 with a standard deviation of $15,000. What is the ***standard deviation*** of the ***total*** amount raised over 10 days? <u>Answer with a number *and* your work. [No partial credit]</u>

**(7)** See the ***Supplement for Question (7)***: *Mortality Inequality.*

**(a)** [3 pts] Consider $\widehat{M}_{it} = 1.31 + 0.015\ PP_{it} - 0.45\ YR2010_{it} - 0.009\ YR2010_{it} * PP_{it}$ , where $\widehat{M}_{it}$ is the predicted three-year mortality rate per 1,000 population, $PP_{it}$ is the poverty percentile, $YR2010_{it}$ is a dummy variable equal to 1 for the year 2010 and 0 otherwise, and $i$ indexes the county groups and $t$ the year. Which age range (0-4, 5-19, 20-49, or 50+) and which sex (female or male) does the given OLS equation correspond to? <u>Answer with the age range and sex. [No partial credit]</u>

**(b)** [3 pts] For the same age range and sex as Part (a), find $b_0, b_1, b_2$, and $b_3$ in an alternative specification: $\widehat{M}_{it} = b_0 + b_1\ PP_{it} + b_2\ YR1990_{it} + b_3\ YR1990_{it} * PP_{it}$ , where $YR1990_{it}$ is a dummy variable equal to 1 for the year 1990. <u>Answer with an equation with numbers replacing $b_0, b_1, b_2$, and $b_3$. [No partial credit]</u>

**(c)** [10 pts] Look carefully at the figure for ***females 50 years old and older***. In the appendix, the authors report the slopes: for 1990 it is 0.098 and for 2010 it is 0.158. They also report a P-value of 0.032 for a test of a difference in slopes. For females 50 years old and older, what can we conclude about changes in mortality rates and changes in mortality inequality between 1990 and 2010? Also, what makes this group different from the other seven groups (i.e. males of all ages and females below 50)? Explain clearly, accurately and at a level that a reader of *The Economist* could understand (plain English). <u>Answer with 4 – 6 sentences.</u>

**(8)** See the ***Supplement for Question (8)****: Risky Mortgage Securities and the 2008 Financial Crisis.*

**(a)** [4 pts] Consider the "perfectly uncorrelated" scenario and two events defined as: "the first mortgage defaults" and "the second mortgage defaults." Are these two events mutually exclusive (i.e. disjoint)? Are these two events independent? Are these two events both mutually exclusive and independent? Answer with 2 – 3 sentences.

**(b)** [9 pts] Show the work for computing *each of the five probabilities* listed under the column entitled "If defaults are perfectly uncorrelated" in Table 1. Answer with a quantitative analysis.

This *Supplement* contains graphs, tables, and other information needed to answer some of the exam questions as well as the aid sheets (formula sheets and Normal, $t$ and $F$ statistical tables). For each question directing you to this *Supplement*, carefully review all relevant materials. Remember, *only* your answers written on the exam papers (in the designated space immediately after each question) will be graded. Any writing on this *Supplement* will *not* be graded.

***Supplement for Question (2):*** *The Economist* produces "The Big Mac Index" (http://www.economist.com/content/big-mac-index) using data from various sources. Part of those data[1] are reproduced to the right, showing the price of a Big Mac hamburger in China between June 2005 and January 2017. For each reported time period, it shows the local price, which is in Chinese Yuan ¥.

The Excel regression output, given below, shows a regression of the local price on the time period measured by t, which is the number of months since June 2005.

| year | month | t | country | local_price |
|---|---|---|---|---|
| 2005 | June | 0 | China | 10.50 |
| 2006 | January | 7 | China | 10.50 |
| 2006 | May | 11 | China | 10.50 |
| 2007 | January | 19 | China | 11.00 |
| 2007 | June | 24 | China | 11.00 |
| 2008 | June | 36 | China | 12.50 |
| 2009 | July | 49 | China | 12.50 |
| 2010 | January | 55 | China | 12.50 |
| 2010 | July | 61 | China | 13.20 |
| 2011 | July | 73 | China | 14.65 |
| 2012 | January | 79 | China | 15.40 |
| 2012 | July | 85 | China | 15.65 |
| 2013 | January | 91 | China | 16.00 |
| 2013 | July | 97 | China | 16.00 |
| 2014 | January | 103 | China | 16.60 |
| 2014 | July | 109 | China | 16.90 |
| 2015 | January | 115 | China | 17.20 |
| 2015 | July | 121 | China | 17.00 |
| 2016 | January | 127 | China | 17.60 |
| 2016 | July | 133 | China | 18.60 |
| 2017 | January | 139 | China | 19.60 |

**Excel regression output:**

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.988610503 |
| R Square | 0.977350727 |
| Adjusted R Square | 0.97615866 |
| Standard Error | 0.448771056 |
| Observations | 21 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 1 | 165.1199148 | 165.1199148 | 819.8790298 | 4.30887E-17 |
| Residual | 19 | 3.826513751 | 0.201395461 | | |
| Total | 20 | 168.9464286 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
|---|---|---|---|---|---|---|
| Intercept | 9.818475368 | 0.191868507 | 51.17293888 | 8.03224E-22 | 9.416889968 | 10.22006077 |
| t | 0.06467537 | 0.002258728 | 28.63352982 | 4.30887E-17 | 0.059947797 | 0.069402943 |

[1] http://infographics.economist.com/2017/databank/BMFile2000toJan2017.xls, retrieved on February 17, 2017.

***Supplement for Question (3):*** Consider the 2013 paper "Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant" from the *Journal of Political Economy* (DOI: 10.1086/671137). They use extensive data at the daily level for an automobile manufacturer that started production of completely redesigned vehicles.

**EXCERPT (pp. 653 - 654):** Figure 1 plots the average number of defects per car by week. When production begins in mid-August, average defect rates were around 75 per car. Eight weeks later, they had fallen by two-thirds, to roughly 25 defects per car. These strong initial learning effects are consistent with findings in the broader literature on learning by doing.

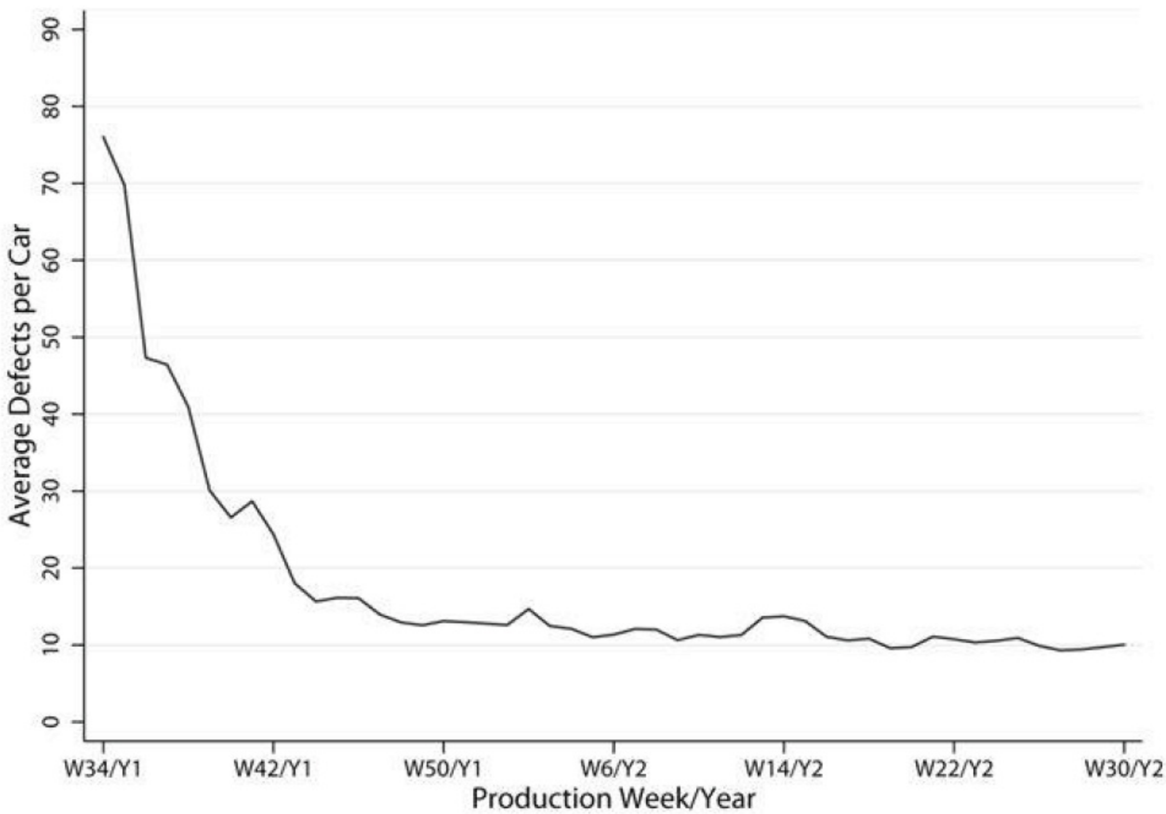



**Figure 1:** Average defect rates per car. The figure plots the average number of production defects per car by week over the production year. Weeks are labeled on the horizontal axis; for example, W34/Y1 indicates the thirty-fourth calendar week of calendar year 1 (the production spanned two calendar years, from August of year 1 to June of year 2).

The researchers start with this simple empirical model of the learning process:

$$\ln(D_t) = \alpha + \beta \ln(E_t) + \varepsilon_t$$

where $t$ indexes either a day or a week, $D_t$ is the average defects per car in a time period, and $E_t$ is the production experience up to that point (cumulative production). The researchers use the natural logarithm (which they abbreviate either as log or ln) to transform the original data. They also explore an alternative model that includes a time trend:

$$\ln(D_t) = \alpha + \beta \ln(E_t) + \gamma * t + \varepsilon_t$$

where $t$ is a variable measuring the number of time periods since the start of production.

**EXCERPT (p. 655):** Table 1 shows the results of estimating these specifications with our sample. Panel A contains the results from specifications using weekly data (average defect rates over the week and production experience at the week's outset); Panel B shows results obtained using daily observations. [Next is an excerpt of Table 1, slightly modified for clarity.]

***Supplement for Question (3) continues on next page >>>>>***

***Supplement for Question (3), cont'd:***

**Table 1:** Estimates of Learning By Doing

| | (1) | (2) |
|---|---|---|
| | Panel A. Weekly Data | |
| Estimated learning rate, $\hat{\beta}$ | -0.289*<br>(0.007) | -0.336*<br>(0.017) |
| Time trend | | 0.007*<br>(0.002) |
| Observations | 47 | 47 |
| $R^2$ | 0.961 | 0.969 |
| | Panel B. Daily Data | |
| Estimated learning rate, $\hat{\beta}$ | -0.306*<br>(0.006) | -0.369*<br>(0.014) |
| Time trend | | 0.001*<br>(0.0002) |
| Observations | 224 | 224 |
| $R^2$ | 0.931 | 0.943 |

*Notes:* Column (1) in both panels shows estimation results for $\ln(D_t) = \alpha + \beta \ln(E_t) + \varepsilon_t$ , where $D_t$ is the average defects per car in time period $t$ and $E_t$ is production experience up to that point: cumulative number of cars produced before the current period. Column (2) in both panels shows estimation results for $\ln(D_t) = \alpha + \beta \ln(E_t) + \gamma * t + \varepsilon_t$. Heteroskedasticity-robust standard errors are in parentheses. * Significant at the 5 percent level.

**EXCERPT (p. 656):** The simple empirical model fits the data very well at both frequencies, with the $R^2$ of the weekly and daily specifications at 0.961 and 0.931, respectively. This fit can also be seen in Figure 2, which plots the logged average defect rate against cumulative production in the daily data.

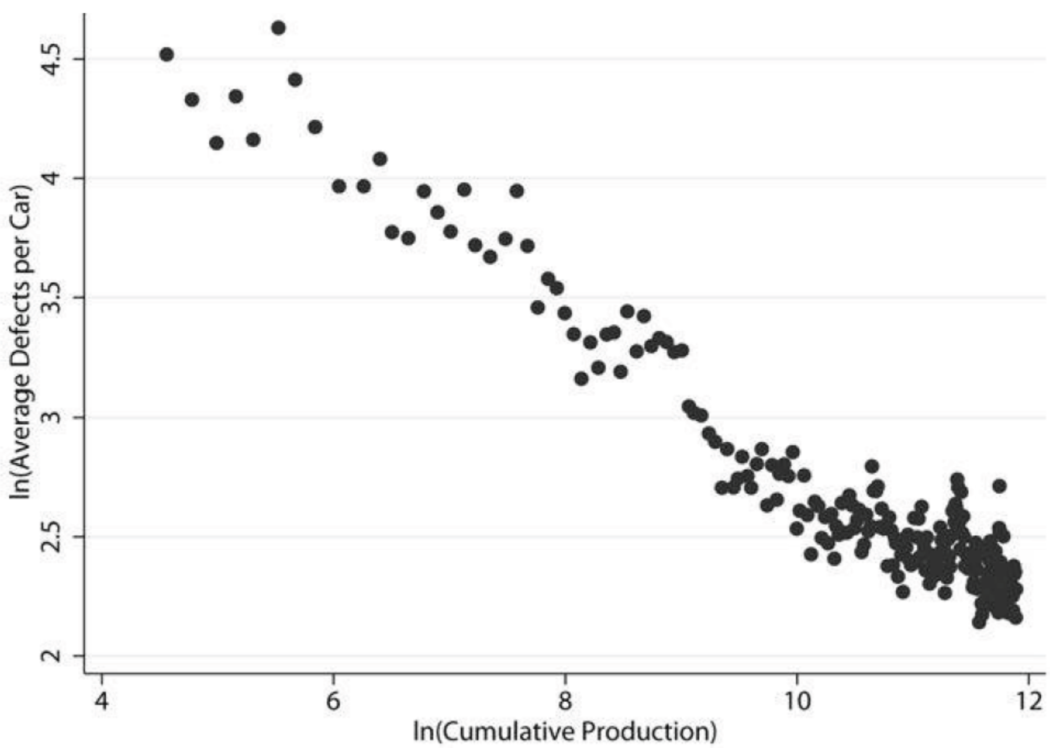



**Figure 2:** Log defects per car versus log production experience (cumulative output), daily data. The figure plots daily data on the (logged) average number of production defects per car versus (logged) cumulative production. Cumulative production is the cumulative number of cars produced before the day of observation.

***Supplement for Question (4):*** Consider a 2016 *NBER* Working Paper "A New Approach to an Age-Old Problem: Solving Externalities by Incenting Workers Directly," updated in 2017 (http://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2017/01/Working-paper-262-Gosnell-et-al.pdf). The authors partnered with Virgin Atlantic Airways to conduct an experiment with their captains and their co-pilots. The goal is to identify effective ways to get captains to improve their fuel efficiency. There are three stages of a flight: (1) pre-flight, (2) in-flight, and (3) post-flight. Captains make choices that affect fuel efficiency in all three stages. Prior to working with these researchers, the airline, like others in the industry, did not give its captains and co-pilots any feedback on how fuel-efficient they were. In the field experiment, each of the 335 Virgin Atlantic Airways captains is *randomly assigned* to one of four groups:

- Control Group ("Control"): Each captain continued not to receive any feedback about her/his fuel-efficiency, but is informed that fuel-efficiency is now being monitored for all pilots.
- Treatment Group 1 ("Information"): At the end of each month each captain receives a report saying which percent of her/his flights met the fuel-efficiency targets.
- Treatment Group 2 ("Targets"): In addition to monthly reports, each captain gets personalized targets (based on fuel-efficiency before the experiment) and monthly feedback on how well s/he is meeting them.
- Treatment Group 3 ("Prosocial"): In addition to monthly reports and personalized targets, each captain also raises money for her/his favorite charity by meeting her/his personalized targets.

The researchers conclude that all four methods are highly effective in improving fuel efficiency. Further, because these methods are low-cost, they conclude that targeting individual workers is important for reducing firms' negative externalities and for fighting climate change.

We focus on *one* of the three stages of a flight and an *in-flight* efficiency variable (which they call a "metric").

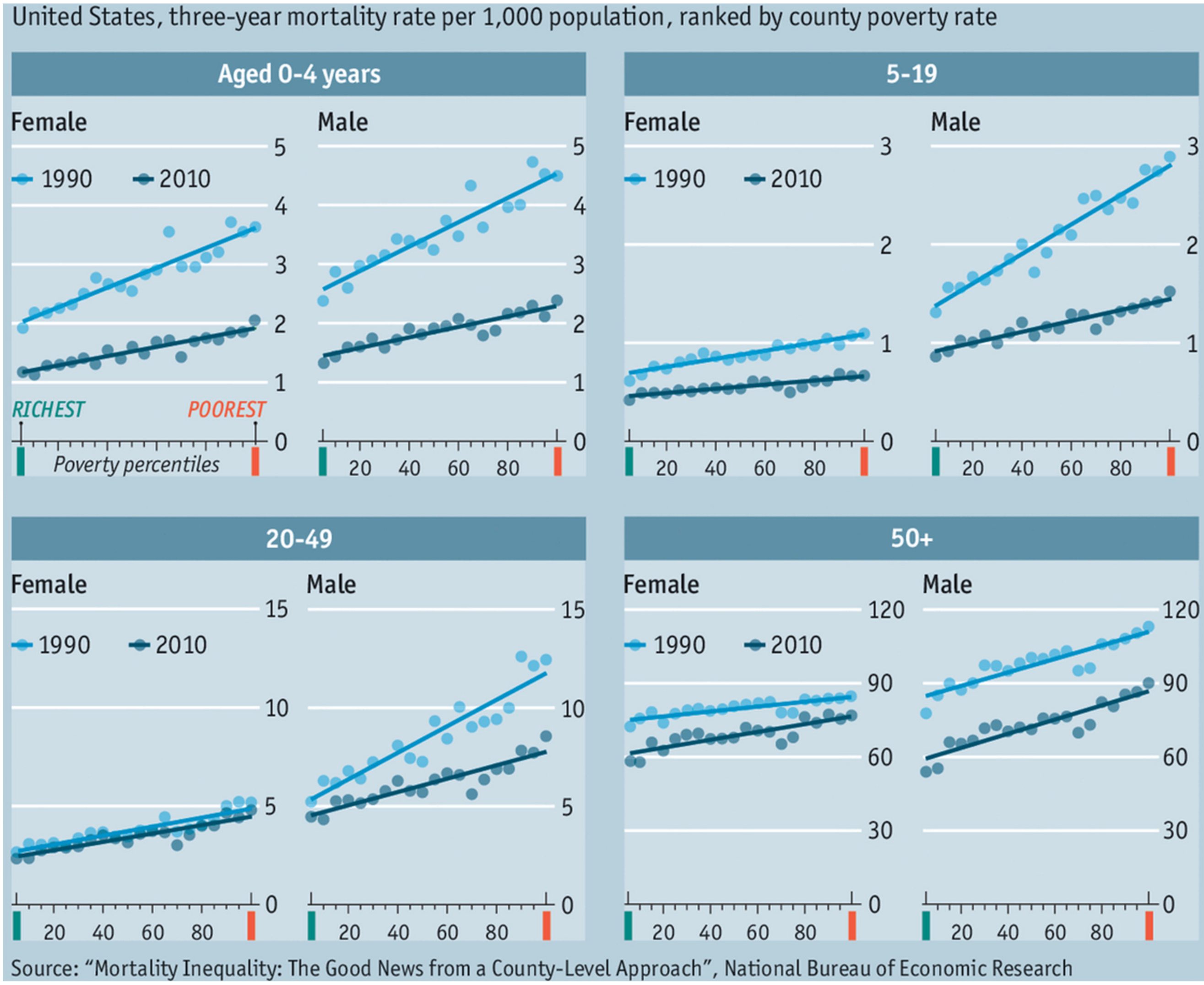
> **EXCERPT (pp. 8 – 9):** The *Efficient Flight* metric captures whether captains use less fuel during flight than is allotted in the flight plan. We use this metric to understand whether captains have made fuel-efficient choices between takeoff and landing. This measure incorporates several in-flight behaviors, such as requesting and executing optimal altitudes and shortcuts from air traffic control, maintaining ideal speeds, and optimally adjusting to en route weather updates. [Sometimes] captains sacrifice fuel efficiency [for safety reasons], so we would not expect even a "model" captain to perform this metric on 100% of flights. In our analysis, *Efficient Flight* equals 1 if the captain does not exceed the projected fuel use for that flight, and 0 otherwise.

**Table 4:** Proportion of Efficient Flights by Group and by Time Period

| | Control | Treatment 1 "Information" | Treatment 2 "Targets" | Treatment 3 "Prosocial" | All Captains |
|---|---|---|---|---|---|
| Before Experiment | 0.311<br>5,258 obs. | 0.314<br>5,429 obs. | 0.313<br>5,070 obs. | 0.312<br>5,140 obs. | 0.312<br>20,897 obs. |
| During Experiment | 0.476<br>3,321 obs. | 0.503<br>3,330 obs. | 0.528<br>3,016 obs. | 0.510<br>3,258 obs. | 0.504<br>12,925 obs. |
| After Experiment | 0.548<br>2,140 obs. | 0.521<br>2,120 obs. | 0.536<br>1,867 obs. | 0.525<br>2,063 obs. | 0.533<br>8,190 obs. |

*Notes:* The table reports the proportion of flights for which captains in a given group achieved an *Efficient Flight*. Also reported are the number of observations (flights) from which the summary statistics are calculated.

***Supplement for Question (7):*** On May 14, 2016 *The Economist* posted “Looking up: The link between income and mortality rates is weakening.” It includes the figure below with the title “Lower, flatter, better.”




The y-axis is the three-year mortality rate per 1,000 population.[2] The source is a 2016 article “Mortality Inequality: The Good News from a County-Level Approach” in the *Journal of Economic Perspectives* (doi=10.1257/jep.30.2.29). The excerpt shows how the economists who conducted the original research interpreted the top half of the figure.

> **EXCERPT (p. 40):** [The figure] shows three-year mortality rates at the level of county groups, with counties ranked by the share of their population below the poverty line, for males and females in four different age groups. In these figures, each marker shows the mortality rate for a bin representing 5 percent of the US population in the relevant year. A slope that becomes steeper over time implies increasing inequality and vice versa.
>
> [The figure] shows dramatic reductions in mortality among children aged zero to four between 1990 and 2010. Overall, the reductions in under-five mortality were much greater in poorer counties than in richer ones, and slightly larger for males than for females. For example, the under-five mortality rate for males fell from 4.5 per 1,000 in 1990 to 2.4 per 1,000 in the poorest counties, compared to a decline from 2.4 to 1.3 per 1,000 in the richest counties over the same period. Among children aged 5 to 19, there were large reductions in mortality for males, with more modest reductions for females (from already low levels). Once again, reductions were larger in poorer counties, implying significant reductions in mortality inequality.

---

[2] Page 37 explains: “the three-year mortality rate in 1990 is the ratio of all deaths that occurred in a cohort [sex and age group] between April 1, 1990, and March 31, 1993, divided by the 1990 Census population count [for that sex and age group].”

***Supplement for Question (8):*** Consider a 2012 book *The Signal and the Noise: Why So Many Predictions Fail – but Some Don't* by Nate Silver. The excerpt below is taken from a chapter on the 2008 financial crisis and how major rating agencies gave AAA ratings, signalling the safest investments, to risky and complicated securities called collateralized debt obligations (CDOs). The excerpt has been abbreviated and clarified in places. Also, to assist your reading, the key sentences have been underlined. Make sure to note the underlined sentences and Table 1 below.

**EXCERPT (pp. 26 – 28):** CDOs are collections of mortgage debt. Imagine you have a set of five mortgages, each of which you assume has a 5 percent chance of defaulting. You can create a number of bets based on the status of these mortgages, each of which is progressively more risky.

The safest of these bets, what I'll call the Alpha Pool, pays out unless *all five* of the mortgages default. The riskiest, the Epsilon Pool, leaves you on the hook if *any* of the five mortgages defaults. Then there are other steps along the way.

Why might an investor prefer making a bet on the Epsilon Pool to the Alpha Pool? That's easy – because it will be priced more cheaply to account for the greater risk. But say you're a risk-averse investor, such as a pension fund. If you're going to buy anything, it will be the Alpha Pool, which will assuredly be rated AAA.

The Alpha Pool consists of five mortgages, each of which has only a 5 percent chance of defaulting. You lose the bet only if *all five* actually do default. What is the risk of that happening?

Actually, that is not an easy question – and therein lies the problem. The assumptions you choose will yield profoundly different answers. If you make the wrong assumptions, your model may be extraordinarily wrong.

One assumption is that each mortgage is independent of the others. In this scenario, your risks are well diversified: if a carpenter in Cleveland defaults on his mortgage, this will have no bearing on whether a dentist in Denver does. Under this scenario, the risk of losing your bet would be exceptionally small (0.00003%). This supposed miracle of diversification is how the rating agencies claimed that a group of subprime mortgages that had just a B+ credit rating on average – which would ordinarily imply more than a 20 percent chance of default – had almost no chance of defaulting when pooled together.

The other extreme is to assume that the mortgages, instead of being entirely independent of one another, will all behave exactly alike. That is, either all five mortgages will default or none will. There's a 5 percent chance that all the mortgages will default – making your bet *160,000 times riskier* than you had thought originally.

**Table 1:** Simplified CDO Structure

| | | **PROBABILITY OF LOSING BET** | | |
|---|---|---|---|---|
| **Bet** | **Rules** | **If defaults are perfectly uncorrelated** | **If defaults are perfectly correlated** | **Risk multiple** |
| Alpha Pool | Bet wins unless *all* 5 mortgages default | 0.00003% | 5.0% | 160,000x |
| Beta Pool | Bet wins unless *exactly* 4 of 5 mortgages default | 0.003% | 5.0% | 1,684x |
| Gamma Pool | Bet wins unless *exactly* 3 of 5 mortgages default | 0.1% | 5.0% | 44x |
| Delta Pool | Bet wins unless *exactly* 2 of 5 mortgages default | 2.1% | 5.0% | 2.3x |
| Epsilon Pool | Bet wins unless *any* of 5 mortgages default | 22.6% | 5.0% | 0.2x |

Which of these assumptions is more valid will depend on economic conditions. If the economy and the housing market are healthy, the first scenario – the five mortgages have nothing to do with one another – might be a reasonable approximation. Defaults are going to happen from time to time because of unfortunate events: someone gets hit with a huge medical bill, or they lose their job. However, one person's default risk won't have much to do with another's.

But suppose instead that there is some common factor that ties the fate of these homeowners together. For instance: there is a massive housing bubble that has caused home prices to rise by 80 percent without any tangible improvement in fundamentals. Now you've got trouble: if one borrower defaults, the rest might succumb to the same problem. The risk of losing your bet has increased by orders of magnitude.

**Sample mean:** $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ **Sample variance:** $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2}{n-1} - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$ **Sample s.d.:** $s = \sqrt{s^2}$

**Sample coefficient of variation:** $CV = \frac{s}{\bar{X}}$ **Sample covariance:** $s_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y})}{n-1} = \frac{\sum_{i=1}^{n} x_i y_i}{n-1} - \frac{(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n(n-1)}$

**Sample interquartile range:** $IQR = Q3 - Q1$ **Sample coefficient of correlation:** $r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n} z_{x_i} z_{y_i}}{n-1}$

**Addition rule:** $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$ **Conditional probability:** $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$

**Complement rules:** $P(A^C) = P(A') = 1 - P(A)$ $P(A^C|B) = P(A'|B) = 1 - P(A|B)$

**Multiplication rule:** $P(A \text{ and } B) = P(A|B)P(B) = P(B|A)P(A)$

**Expected value:** $E[X] = \mu = \sum_{all\ x} xp(x)$ **Variance:** $V[X] = E[(X-\mu)^2] = \sigma^2 = \sum_{all\ x}(x-\mu)^2 p(x)$

**Covariance:** $COV[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY} = \sum_{all\ x}\sum_{all\ y}(x-\mu_X)(y-\mu_Y)p(x,y)$

**<u>Laws of expected value</u>:**

$E[c] = c$

$E[X+c] = E[X] + c$

$E[cX] = cE[X]$

$E[a+bX+cY] = a + bE[X] + cE[Y]$

**<u>Laws of variance</u>:**

$V[c] = 0$

$V[X+c] = V[X]$

$V[cX] = c^2 V[X]$

$V[a+bX+cY] = b^2V[X] + c^2V[Y] + 2bc * COV[X,Y]$

$V[a+bX+cY] = b^2V[X] + c^2V[Y] + 2bc * SD(X) * SD(Y) * \rho$

where $\rho = CORRELATION[X,Y]$

**<u>Laws of covariance</u>:**

$COV[X,c] = 0$

$COV[a+bX, c+dY] = bd * COV[X,Y]$

**Combinatorial formula:** $C_x^n = \frac{n!}{x!(n-x)!}$ **Binomial probability:** $p(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$ for $x = 0,1,2,\dots,n$

**If *X* is Binomial** $(X \sim B(n,p))$ **then** $E[X] = np$ **and** $V[X] = np(1-p)$

**If *X* is Uniform** $(X \sim U[a,b])$ **then** $f(x) = \frac{1}{b-a}$ **and** $E[X] = \frac{a+b}{2}$ **and** $V[X] = \frac{(b-a)^2}{12}$

**<u>Sampling distribution of $\bar{X}$</u>:**

$\mu_{\bar{X}} = E[\bar{X}] = \mu$

$\sigma_{\bar{X}}^2 = V[\bar{X}] = \frac{\sigma^2}{n}$

$\sigma_{\bar{X}} = SD[\bar{X}] = \frac{\sigma}{\sqrt{n}}$

**<u>Sampling distribution of $\hat{P}$</u>:**

$\mu_{\hat{P}} = E[\hat{P}] = p$

$\sigma_{\hat{P}}^2 = V[\hat{P}] = \frac{p(1-p)}{n}$

$\sigma_{\hat{P}} = SD[\hat{P}] = \sqrt{\frac{p(1-p)}{n}}$

**<u>Sampling distribution of $(\hat{P}_2 - \hat{P}_1)$</u>:**

$\mu_{\hat{P}_2-\hat{P}_1} = E[\hat{P}_2 - \hat{P}_1] = p_2 - p_1$

$\sigma_{\hat{P}_2-\hat{P}_1}^2 = V[\hat{P}_2 - \hat{P}_1] = \frac{p_2(1-p_2)}{n_2} + \frac{p_1(1-p_1)}{n_1}$

$\sigma_{\hat{P}_2-\hat{P}_1} = SD[\hat{P}_2 - \hat{P}_1] = \sqrt{\frac{p_2(1-p_2)}{n_2} + \frac{p_1(1-p_1)}{n_1}}$

**<u>Sampling distribution of $(\bar{X}_1 - \bar{X}_2)$, independent samples</u>:**

$\mu_{\bar{X}_1-\bar{X}_2} = E[\bar{X}_1 - \bar{X}_2] = \mu_1 - \mu_2$

$\sigma_{\bar{X}_1-\bar{X}_2}^2 = V[\bar{X}_1 - \bar{X}_2] = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

$\sigma_{\bar{X}_1-\bar{X}_2} = SD[\bar{X}_1 - \bar{X}_2] = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

**<u>Sampling distribution of $(\bar{X}_d)$, paired ($d = X_1 - X_2$)</u>:**

$\mu_{\bar{X}_d} = E[\bar{X}_d] = \mu_1 - \mu_2$

$\sigma_{\bar{X}_d}^2 = V[\bar{X}_d] = \frac{\sigma_d^2}{n} = \frac{\sigma_1^2 + \sigma_2^2 - 2*\rho*\sigma_1*\sigma_2}{n}$

$\sigma_{\bar{X}_d} = SD[\bar{X}_d] = \frac{\sigma_d}{\sqrt{n}} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2*\rho*\sigma_1*\sigma_2}{n}}$

**Inference about a population proportion:**

***z* test statistic:** $z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ **CI estimator:** $\hat{P} \pm z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

**Inference about comparing two population proportions:**

***z* test statistic under Null hypothesis of no difference:** $z = \frac{\hat{P}_2 - \hat{P}_1}{\sqrt{\frac{\bar{P}(1-\bar{P})}{n_1} + \frac{\bar{P}(1-\bar{P})}{n_2}}}$ **Pooled proportion:** $\bar{P} = \frac{X_1 + X_2}{n_1 + n_2}$

**CI estimator:** $(\hat{P}_2 - \hat{P}_1) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_2(1-\hat{P}_2)}{n_2} + \frac{\hat{P}_1(1-\hat{P}_1)}{n_1}}$

**Inference about the population mean:**

***t* test statistic:** $t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ **CI estimator:** $\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$ **Degrees of freedom:** $\nu = n - 1$

**Inference about a comparing two population means, independent samples, unequal variances:**

***t* test statistic:** $t = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ **CI estimator:** $(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

**Degrees of freedom:** $\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$

**Inference about a comparing two population means, independent samples, assuming equal variances:**

***t* test statistic:** $t = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$ **CI estimator:** $(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$ **Degrees of freedom:** $\nu = n_1 + n_2 - 2$

**Pooled variance:** $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

**Inference about a comparing two population means, paired data:** ($n$ is number of pairs and $d = X_1 - X_2$)

***t* test statistic:** $t = \frac{\bar{d} - \Delta_0}{s_d/\sqrt{n}}$ **CI estimator:** $\bar{X}_d \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}}$ **Degrees of freedom:** $\nu = n - 1$

**SIMPLE REGRESSION:**

**Model:** $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ **OLS line:** $\hat{y}_i = b_0 + b_1 x_i$ $b_1 = \frac{s_{xy}}{s_x^2} = r\frac{s_y}{s_x}$ $b_0 = \bar{Y} - b_1\bar{X}$

**Coefficient of determination:** $R^2 = (r)^2$ **Residuals:** $e_i = y_i - \hat{y}_i$

**Standard deviation of residuals:** $s_e = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n}(e_i - 0)^2}{n-2}}$ **Standard error of slope:** $s.e.(b_1) = s_{b_1} = \frac{s_e}{\sqrt{(n-1)s_x^2}}$

**Inference about the population slope:**

***t* test statistic:** $t = \frac{b_1 - \beta_{10}}{s.e.(b_1)}$ **CI estimator:** $b_1 \pm t_{\alpha/2} s.e.(b_1)$ **Degrees of freedom:** $\nu = n - 2$

**Standard error of slope:** $s.e.(b_1) = s_{b_1} = \frac{s_e}{\sqrt{(n-1)s_x^2}}$

**Prediction interval for $y$ at given value of $x$ ($x_g$):**

$$\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}} \quad \textbf{or} \quad \hat{y}_{x_g} \pm t_{\alpha/2} \sqrt{\left(s.e.(b_1)\right)^2 \left(x_g - \bar{X}\right)^2 + \frac{s_e^2}{n} + s_e^2}$$

**Degrees of freedom:** $\nu = n - 2$

**Confidence interval for predicted mean at given value of $x$ ($x_g$):**

$$\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}} \quad \textbf{or} \quad \hat{y}_{x_g} \pm t_{\alpha/2} \sqrt{\left(s.e.(b_1)\right)^2 \left(x_g - \bar{X}\right)^2 + \frac{s_e^2}{n}}$$

**Degrees of freedom:** $\nu = n - 2$

**SIMPLE & MULTIPLE REGRESSION:**

**Model:** $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i$

$$SST = \sum_{i=1}^{n} (y_i - \bar{Y})^2 = SSR + SSE \quad SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{Y})^2 \quad SSE = \sum_{i=1}^{n} {e_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$s_y^2 = \frac{SST}{n-1} \quad MSE = \frac{SSE}{n-k-1} \quad Root\ MSE = \sqrt{\frac{SSE}{n-k-1}} \quad MSR = \frac{SSR}{k}$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad Adj. R^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} = \left(R^2 - \frac{k}{n-1}\right)\left(\frac{n-1}{n-k-1}\right)$$

**Residuals:** $e_i = y_i - \hat{y}_i$ **Standard deviation of residuals:** $s_e = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{\frac{\sum_{i=1}^{n}(e_i - 0)^2}{n-k-1}}$

**Inference about the overall statistical significance of the regression model:**

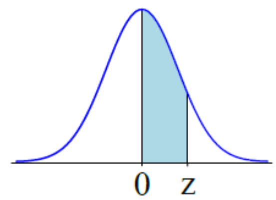
$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{(SST-SSE)/k}{SSE/(n-k-1)} = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE}$$

**Numerator degrees of freedom:** $\nu_1 = k$ **Denominator degrees of freedom:** $\nu_2 = n - k - 1$

**Inference about the population slope for explanatory variable j:**

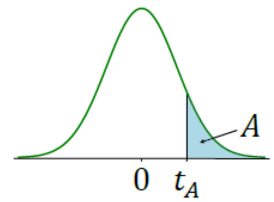
***t* test statistic:** $t = \frac{b_j - \beta_{j0}}{s_{b_j}}$ **CI estimator:** $b_j \pm t_{\alpha/2} s_{b_j}$ **Degrees of freedom:** $\nu = n - k - 1$

**Standard error of slope:** $s.e.(b_j) = s_{b_j}$ (for multiple regression, must be obtained from technology)




**Normal Probabilities:**

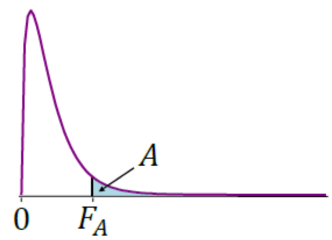
| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1 | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2 | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3 | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.1480 | 0.1517 |
| 0.4 | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5 | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.2190 | 0.2224 |
| 0.6 | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.2486 | 0.2517 | 0.2549 |
| 0.7 | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.2794 | 0.2823 | 0.2852 |
| 0.8 | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.3078 | 0.3106 | 0.3133 |
| 0.9 | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.3340 | 0.3365 | 0.3389 |
| 1.0 | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.3577 | 0.3599 | 0.3621 |
| 1.1 | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.3770 | 0.3790 | 0.3810 | 0.3830 |
| 1.2 | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.3980 | 0.3997 | 0.4015 |
| 1.3 | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |
| 1.4 | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.4292 | 0.4306 | 0.4319 |
| 1.5 | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 | 0.4394 | 0.4406 | 0.4418 | 0.4429 | 0.4441 |
| 1.6 | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.4525 | 0.4535 | 0.4545 |
| 1.7 | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.4616 | 0.4625 | 0.4633 |
| 1.8 | 0.4641 | 0.4649 | 0.4656 | 0.4664 | 0.4671 | 0.4678 | 0.4686 | 0.4693 | 0.4699 | 0.4706 |
| 1.9 | 0.4713 | 0.4719 | 0.4726 | 0.4732 | 0.4738 | 0.4744 | 0.4750 | 0.4756 | 0.4761 | 0.4767 |
| 2.0 | 0.4772 | 0.4778 | 0.4783 | 0.4788 | 0.4793 | 0.4798 | 0.4803 | 0.4808 | 0.4812 | 0.4817 |
| 2.1 | 0.4821 | 0.4826 | 0.4830 | 0.4834 | 0.4838 | 0.4842 | 0.4846 | 0.4850 | 0.4854 | 0.4857 |
| 2.2 | 0.4861 | 0.4864 | 0.4868 | 0.4871 | 0.4875 | 0.4878 | 0.4881 | 0.4884 | 0.4887 | 0.4890 |
| 2.3 | 0.4893 | 0.4896 | 0.4898 | 0.4901 | 0.4904 | 0.4906 | 0.4909 | 0.4911 | 0.4913 | 0.4916 |
| 2.4 | 0.4918 | 0.4920 | 0.4922 | 0.4925 | 0.4927 | 0.4929 | 0.4931 | 0.4932 | 0.4934 | 0.4936 |
| 2.5 | 0.4938 | 0.4940 | 0.4941 | 0.4943 | 0.4945 | 0.4946 | 0.4948 | 0.4949 | 0.4951 | 0.4952 |
| 2.6 | 0.4953 | 0.4955 | 0.4956 | 0.4957 | 0.4959 | 0.4960 | 0.4961 | 0.4962 | 0.4963 | 0.4964 |
| 2.7 | 0.4965 | 0.4966 | 0.4967 | 0.4968 | 0.4969 | 0.4970 | 0.4971 | 0.4972 | 0.4973 | 0.4974 |
| 2.8 | 0.4974 | 0.4975 | 0.4976 | 0.4977 | 0.4977 | 0.4978 | 0.4979 | 0.4979 | 0.4980 | 0.4981 |
| 2.9 | 0.4981 | 0.4982 | 0.4982 | 0.4983 | 0.4984 | 0.4984 | 0.4985 | 0.4985 | 0.4986 | 0.4986 |
| 3.0 | 0.4987 | 0.4987 | 0.4987 | 0.4988 | 0.4988 | 0.4989 | 0.4989 | 0.4989 | 0.4990 | 0.4990 |
| 3.1 | 0.4990 | 0.4991 | 0.4991 | 0.4991 | 0.4992 | 0.4992 | 0.4992 | 0.4992 | 0.4993 | 0.4993 |
| 3.2 | 0.4993 | 0.4993 | 0.4994 | 0.4994 | 0.4994 | 0.4994 | 0.4994 | 0.4995 | 0.4995 | 0.4995 |
| 3.3 | 0.4995 | 0.4995 | 0.4995 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4997 |
| 3.4 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4998 |
| 3.5 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 |
| 3.6 | 0.4998 | 0.4998 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 |




**Critical Values of $t$:**

| $\nu$ | $t_{0.10}$ | $t_{0.05}$ | $t_{0.025}$ | $t_{0.01}$ | $t_{0.005}$ | $\nu$ | $t_{0.10}$ | $t_{0.05}$ | $t_{0.025}$ | $t_{0.01}$ | $t_{0.005}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 | 38 | 1.304 | 1.686 | 2.024 | 2.429 | 2.712 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 39 | 1.304 | 1.685 | 2.023 | 2.426 | 2.708 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 41 | 1.303 | 1.683 | 2.020 | 2.421 | 2.701 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 42 | 1.302 | 1.682 | 2.018 | 2.418 | 2.698 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 43 | 1.302 | 1.681 | 2.017 | 2.416 | 2.695 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 44 | 1.301 | 1.680 | 2.015 | 2.414 | 2.692 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 45 | 1.301 | 1.679 | 2.014 | 2.412 | 2.690 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 46 | 1.300 | 1.679 | 2.013 | 2.410 | 2.687 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 47 | 1.300 | 1.678 | 2.012 | 2.408 | 2.685 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 48 | 1.299 | 1.677 | 2.011 | 2.407 | 2.682 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 49 | 1.299 | 1.677 | 2.010 | 2.405 | 2.680 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 50 | 1.299 | 1.676 | 2.009 | 2.403 | 2.678 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 51 | 1.298 | 1.675 | 2.008 | 2.402 | 2.676 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 52 | 1.298 | 1.675 | 2.007 | 2.400 | 2.674 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 53 | 1.298 | 1.674 | 2.006 | 2.399 | 2.672 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 54 | 1.297 | 1.674 | 2.005 | 2.397 | 2.670 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 55 | 1.297 | 1.673 | 2.004 | 2.396 | 2.668 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 65 | 1.295 | 1.669 | 1.997 | 2.385 | 2.654 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 70 | 1.294 | 1.667 | 1.994 | 2.381 | 2.648 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 75 | 1.293 | 1.665 | 1.992 | 2.377 | 2.643 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 80 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 90 | 1.291 | 1.662 | 1.987 | 2.368 | 2.632 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 100 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 120 | 1.289 | 1.658 | 1.980 | 2.358 | 2.617 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 140 | 1.288 | 1.656 | 1.977 | 2.353 | 2.611 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 160 | 1.287 | 1.654 | 1.975 | 2.350 | 2.607 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 180 | 1.286 | 1.653 | 1.973 | 2.347 | 2.603 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 200 | 1.286 | 1.653 | 1.972 | 2.345 | 2.601 |
| 31 | 1.309 | 1.696 | 2.040 | 2.453 | 2.744 | 250 | 1.285 | 1.651 | 1.969 | 2.341 | 2.596 |
| 32 | 1.309 | 1.694 | 2.037 | 2.449 | 2.738 | 300 | 1.284 | 1.650 | 1.968 | 2.339 | 2.592 |
| 33 | 1.308 | 1.692 | 2.035 | 2.445 | 2.733 | 400 | 1.284 | 1.649 | 1.966 | 2.336 | 2.588 |
| 34 | 1.307 | 1.691 | 2.032 | 2.441 | 2.728 | 500 | 1.283 | 1.648 | 1.965 | 2.334 | 2.586 |
| 35 | 1.306 | 1.690 | 2.030 | 2.438 | 2.724 | 750 | 1.283 | 1.647 | 1.963 | 2.331 | 2.582 |
| 36 | 1.306 | 1.688 | 2.028 | 2.434 | 2.719 | 1000 | 1.282 | 1.646 | 1.962 | 2.330 | 2.581 |
| 37 | 1.305 | 1.687 | 2.026 | 2.431 | 2.715 | $\infty$ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |

Degrees of freedom: $\nu$




**Critical Values of $F$: $A = 0.10$**

| $\nu_2$ \ $\nu_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 4.06 | 3.78 | 3.62 | 3.52 | 3.45 | 3.40 | 3.37 | 3.34 | 3.32 | 3.30 | 3.28 | 3.27 |
| 10 | 3.29 | 2.92 | 2.73 | 2.61 | 2.52 | 2.46 | 2.41 | 2.38 | 2.35 | 2.32 | 2.30 | 2.28 |
| 15 | 3.07 | 2.70 | 2.49 | 2.36 | 2.27 | 2.21 | 2.16 | 2.12 | 2.09 | 2.06 | 2.04 | 2.02 |
| 20 | 2.97 | 2.59 | 2.38 | 2.25 | 2.16 | 2.09 | 2.04 | 2.00 | 1.96 | 1.94 | 1.91 | 1.89 |
| 30 | 2.88 | 2.49 | 2.28 | 2.14 | 2.05 | 1.98 | 1.93 | 1.88 | 1.85 | 1.82 | 1.79 | 1.77 |
| 40 | 2.84 | 2.44 | 2.23 | 2.09 | 2.00 | 1.93 | 1.87 | 1.83 | 1.79 | 1.76 | 1.74 | 1.71 |
| 60 | 2.79 | 2.39 | 2.18 | 2.04 | 1.95 | 1.87 | 1.82 | 1.77 | 1.74 | 1.71 | 1.68 | 1.66 |
| 120 | 2.75 | 2.35 | 2.13 | 1.99 | 1.90 | 1.82 | 1.77 | 1.72 | 1.68 | 1.65 | 1.63 | 1.60 |
| $\infty$ | 2.71 | 2.30 | 2.08 | 1.94 | 1.85 | 1.77 | 1.72 | 1.67 | 1.63 | 1.60 | 1.57 | 1.55 |

**Critical Values of $F$: $A = 0.05$**

| $\nu_2$ \ $\nu_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.70 | 4.68 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.94 | 2.91 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | 2.54 | 2.51 | 2.48 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 | 2.35 | 2.31 | 2.28 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 | 2.16 | 2.13 | 2.09 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.12 | 2.08 | 2.04 | 2.00 |
| 60 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 | 1.99 | 1.95 | 1.92 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.18 | 2.09 | 2.02 | 1.96 | 1.91 | 1.87 | 1.83 |
| $\infty$ | 3.84 | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.88 | 1.83 | 1.79 | 1.75 |

**Critical Values of $F$: $A = 0.01$**

| $\nu_2$ \ $\nu_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.46 | 10.29 | 10.16 | 10.05 | 9.96 | 9.89 |
| 10 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 | 4.85 | 4.77 | 4.71 |
| 15 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.89 | 3.80 | 3.73 | 3.67 |
| 20 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.46 | 3.37 | 3.29 | 3.23 |
| 30 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.30 | 3.17 | 3.07 | 2.98 | 2.91 | 2.84 |
| 40 | 7.31 | 5.18 | 4.31 | 3.83 | 3.51 | 3.29 | 3.12 | 2.99 | 2.89 | 2.80 | 2.73 | 2.66 |
| 60 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.95 | 2.82 | 2.72 | 2.63 | 2.56 | 2.50 |
| 120 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.79 | 2.66 | 2.56 | 2.47 | 2.40 | 2.34 |
| $\infty$ | 6.64 | 4.61 | 3.78 | 3.32 | 3.02 | 2.80 | 2.64 | 2.51 | 2.41 | 2.32 | 2.25 | 2.18 |

Numerator degrees of freedom: $\nu_1$; Denominator degrees of freedom: $\nu_2$