ECO220Y, Term Test #4

March 11, 2016, 9:10 – 11:00 am

U of T e-mail: ___________________________________**@mail.utoronto.ca**

Surname (last name):																			
Given name (first name):																			

UTORID: (e.g. lihao8)								

Instructions:

- You have 110 minutes. Keep these test papers closed on your desk until the start of the test is announced.
- You may use a non-programmable calculator.
- There are 4 questions (some with multiple parts) with varying point values worth a total of 84 points.
- Write your answers clearly, completely and concisely 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*. Your entire answer must fit in the designated space provided immediately after each question.
    - **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.
- Clearly show your work. Make your reasoning clear. **Unless otherwise specified, you choose the significance level. (If there are no special considerations, you may choose a 5% significance level.)**
- Apply your understanding to the specific questions asked. Offer context-specific explanations rather than generic definitions or quotes from class or the book. Show that you can successfully *apply* your understanding to the specific circumstances presented.
- A guide for your response ends each question. The guide lets you know what is expected: e.g. a quantitative analysis, a graph, and/or sentences. If the question and/or guide ask for a fully-labeled graph, it is required.
- For questions with multiple parts (e.g (a) – (c)), attempt each part even if you had trouble with earlier parts.
- This test has 8 pages plus the *Supplement*. The *Supplement* contains the aid sheets (formula sheets, Standard Normal table, and Student t table) as well as graphs, tables, and other information needed to answer the test questions. Anything written on the *Supplement* will *not* be graded. You must write your answers in the designated space provided immediately after each question.

(1) [32 pts] See the *Supplement for Question (1)*.

Page Pts:

(a) [6 pts] For Table 1, which hypothesis test do the t-statistics refer to? The table does not provide all of the information needed to compute the t-statistics: what else would you need to know for each budget? Answer with formal hypotheses, the relevant t-statistic formula, and 1 – 2 sentences.

(b) [6 pts] For Table 2 compared to Table 1, do the bigger (in absolute value) t-statistics mean there are larger differences between males and females in ECO220Y than in the original study? Explain. Answer with 3 – 4 sentences.

(c) [6 pts] For Table 3, find the results associated with a P-value of 0.1938 (in boldface). (Do *not* re-calculate these results.) Which specific hypothesis test does it refer to? How should the results and the P-value of 0.1938 be *interpreted*? Answer with formal hypotheses and 2 – 3 sentences.

Page Pts:

(d) [6 pts] For Table 4, find the results associated with a P-value of 0.0023 (in boldface). (Do *not* re-calculate these results.) Which specific hypothesis test does it refer to? How should the results and the P-value of 0.0023 be *interpreted*? Answer with formal hypotheses and 2 – 3 sentences.

(e) [8 pts] Given Tables 1, 2, 3 and 4, what kinds of information are in Tables 3 and 4 that are not already available in Tables 1 and 2? (In other words, what are the reasons for also presenting Tables 3 and 4?) For Tables 3 and 4 overall, which conclusions should you draw? Answer with 5 – 7 sentences.

Page Pts:

(2) [22 pts] See the *Supplement for Question (2)*.

Page Pts:

(a) [4 pts] What is the value of the R-squared, which has been erased from the STATA output? Offer a *full interpretation*. Answer with a quantitative analysis and 1 sentence.

(b) [8 pts] What are the values of "[95% Conf. Interval]", which have been erased from the STATA output? Offer a *full interpretation*. Answer with a quantitative analysis and 1 – 2 sentences.

(c) [8 pts] Forecast the price of a Big Mac for a particular country that has a GDP of $15,000 USD per capita and give the associated margin of error for your forecast. Answer with a quantitative analysis and 1 sentence.

Page Pts:

(d) [2 pts] How would you check for a violation of the assumption that ε_i is Normally distributed (i.e. Assumption #4), which is relevant to the calculations in the previous part? Answer with 1 sentence.

(3) [16 pts] See the *Supplement for Question (3)*.

Page Pts:

(a) [8 pts] Using Thailand in the 2000's as an illustrative example, *explain* how to construct the data used for Table 1. Also, for the data used in the Table 1 regressions, show the data structure, specifying the variables (columns) and observations (rows). Answer with the data structure and 2 – 3 sentences.

				…	
				…	
…	…	…	…	…	…
				…	

(b) [8 pts] How many regressions are reported in Table 1? How do the x and y variables differ across them? What do the results in PANEL B add beyond PANEL A? How well do the results in PANEL B support the authors' main claim in "Asiaphoria Meets Regression to the Mean"? Explain. Answer with 4 – 5 sentences.

(4) [14 pts] After seeing countless commercials claiming one can get cheaper car insurance from an online company, a local insurance agent was concerned that he may lose some customers. To investigate, he randomly selected profiles (type of car, coverage, driving record, etc.) for 12 of his clients and checked online price quotes for their policies. For those 12 clients, the mean local quote is $842.87 with a s.d. of $230.80 and the mean online quote is $729.56 with a s.d. of $251.99. The coefficient of correlation between the local quotes and online quotes is 0.88. How big is the difference between the online and local quotes? *Fully interpret* your results. Answer with a quantitative analysis (using the best approach to fully address the question asked) and 1 – 2 sentences.

Page Pts:

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

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

Supplement for Question (1): On February 14, 2014, February 6, 2015, and on February 12, 2016 students in ECO220Y engaged in an experiment like the original participants from Andreoni and Vesterlund's "Which is the Fair Sex: Gender Differences in Altruism" published in 2001 and hereafter abbreviated as A&V (2001).

> **Abstract from A&V (2001):** We study gender differences in altruism by examining a modified dictator game with varying incomes and prices. Our results indicate that the question "which is the fair sex?" has a complicated answer—when altruism is expensive, women are kinder, but when it is cheap, men are more altruistic. That is, we find that the male and female "demand curves for altruism" cross, and that men are more responsive to price changes. Furthermore, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be "equalitarians" who prefer to share evenly.

Each participant divided tokens between her/himself and another randomly selected participant in the room (whose identity would never be revealed). Each person made eight allocation decisions – budgets 1 through 8 shown below[1] – where the number of tokens and the point values to each person (self and other) varied. Each point is worth $0.10 to all participants in all cases.

1. Divide 40 tokens: *Hold* _______ @ 1 point each, and *Pass* _______ @ 3 points each.
2. Divide 60 tokens: *Hold* _______ @ 1 point each, and *Pass* _______ @ 2 points each.
3. Divide 75 tokens: *Hold* _______ @ 1 point each, and *Pass* _______ @ 2 points each.
4. Divide 60 tokens: *Hold* _______ @ 1 point each, and *Pass* _______ @ 1 point each.
5. Divide 100 tokens: *Hold* _______ @ 1 point each, and *Pass* _______ @ 1 point each.
6. Divide 60 tokens: *Hold* _______ @ 2 points each, and *Pass* _______ @ 1 point each.
7. Divide 75 tokens: *Hold* _______ @ 2 points each, and *Pass* _______ @ 1 point each.
8. Divide 40 tokens: *Hold* _______ @ 3 points each, and *Pass* _______ @ 1 point each.

We attempted to replicate the original study. One difference is that rather than pay everyone for one randomly selected budget as A&V (2001) did using a research grant, ECO220Y (2014, 2015, 2016) paid randomly selected participants using money students donated right before the session and $20.00 per session (total of $220.00) donated by Prof. Murdock.[2]

A&V (2001) had data from four sessions. These included 70 volunteer undergraduates from intermediate and upper-level economics courses at the University of Wisconsin in 1995 and 72 volunteer undergraduates at Iowa State University in 1997 for a total sample of 142 (95 males and 47 females). ECO220Y (2014, 2015, 2016) used data from eleven sessions spread over February 14, 2014 (three sessions), February 6, 2015 (five sessions), and on February 12, 2016 (three sessions): 868 participated (334 males and 534 females).[3]

Next are four tables of results for the ECO220Y (2014, 2015, 2016) and A&V (2001) data. Two of the tables are in a format like "Table II" in A&V (2001).

[1] Both A&V (2001) and ECO220Y (2014, 2015, 2016) randomized the order the eight allocation choices appeared to each participant.
[2] Nearly all students voluntarily donated $2.00 to a collection jar as suggested by Prof. Murdock.
[3] 903 participated but 35 were discarded for not following experiment instructions: tokens did not add up or sex not indicated.

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

Supplement for Question (1), cont'd:

Table 1: Mean Payoff to Other Party from A&V (2001)

Budget	Token endowment	Income m	p_o/p_s	All subjects (n=142)	Males (n=95)	Females (n=47)	t-stat
1	40	4.00	1/3	3.79	4.18	3.01	1.96
2	60	6.00	1/2	4.03	4.30	3.49	1.48
3	75	7.50	1/2	4.68	5.00	4.03	1.53
4	60	6.00	1	1.54	1.36	1.91	-2.26
5	100	10.00	1	2.52	2.33	2.92	-1.42
6	60	12.00	2	1.42	1.21	1.82	-2.07
7	75	15.00	2	1.71	1.42	2.29	-2.35
8	40	12.00	3	0.89	0.67	1.32	-2.97
Average				2.57	2.56	2.60	-0.24

Table 2: Mean Payoff to Other Party from ECO220Y (2014, 2015 and 2016)

Budget	Token endowment	Income m	p_o/p_s	All subjects (n=868)	Males (n=334)	Females (n=534)	t-stat
1	40	4.00	1/3	4.74	5.39	4.34	3.71
2	60	6.00	1/2	4.83	5.41	4.47	3.66
3	75	7.50	1/2	5.97	6.56	5.60	3.03
4	60	6.00	1	2.17	1.98	2.30	-3.17
5	100	10.00	1	3.36	2.98	3.60	-3.80
6	60	12.00	2	1.91	1.64	2.08	-3.80
7	75	15.00	2	2.32	2.00	2.52	-3.63
8	40	12.00	3	1.18	1.01	1.28	-3.14
Average				3.31	3.37	3.27	1.17

Remember, <u>only</u> your answers written on the test papers (in the designated space immediately after each question) will be graded. Any writing on this *Supplement* will *not* be graded.

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

Supplement for Question (1), cont'd:

Table 3: Proportion Passing Any Money: Comparing A&V (2001) with ECO220Y (2014, 2015 and 2016), by Sex

	Males				*Females*			
	Proportion (s.e.)				Proportion (s.e.)			
Budget	A&V (n=95)	ECO220 (n=334)	Difference (s.e.)	P-value (2-tailed)	A&V (n=47)	ECO220 (n=534)	Difference (s.e.)	P-value (2-tailed)
1	0.71 (0.05)	0.81 (0.02)	-0.10 (0.05)	0.0366	0.77 (0.06)	0.87 (0.01)	-0.11 (0.06)	0.0363
2	**0.75 (0.04)**	**0.81 (0.02)**	**-0.06 (0.05)**	**0.1938**	0.85 (0.05)	0.90 (0.01)	-0.05 (0.05)	0.2832
3	0.73 (0.05)	0.82 (0.02)	-0.09 (0.05)	0.0515	0.81 (0.06)	0.90 (0.01)	-0.09 (0.06)	0.0496
4	0.55 (0.05)	0.69 (0.03)	-0.14 (0.06)	0.0124	0.79 (0.06)	0.87 (0.01)	-0.08 (0.06)	0.1192
5	0.55 (0.05)	0.68 (0.03)	-0.14 (0.06)	0.0146	0.72 (0.07)	0.84 (0.02)	-0.12 (0.07)	0.0320
6	0.48 (0.05)	0.58 (0.03)	-0.10 (0.06)	0.0842	0.68 (0.07)	0.78 (0.02)	-0.10 (0.07)	0.1166
7	0.44 (0.05)	0.58 (0.03)	-0.14 (0.06)	0.0165	0.66 (0.07)	0.80 (0.02)	-0.14 (0.07)	0.0242
8	0.37 (0.05)	0.48 (0.03)	-0.11 (0.06)	0.0498	0.64 (0.07)	0.66 (0.02)	-0.03 (0.07)	0.7126

Table 4: Mean Money Passed Conditional on Passing Any Money for a Given Budget: Comparing A&V (2001) with ECO220Y (2014, 2015 and 2016), by Sex

	Males				Females			
	Mean (s.d.)				Mean (s.d.)			
Budget	A&V	ECO220	Difference (s.e.)	P-value (2-tailed)	A&V	ECO220	Difference (s.e.)	P-value (2-tailed)
1	5.93 (3.86)	6.70 (3.96)	-0.77 (0.53)	0.1499	3.93 (2.60)	4.96 (3.22)	-1.03 (0.46)	0.0294
2	5.75 (3.26)	6.70 (3.49)	-0.95 (0.44)	0.0337	4.10 (2.36)	4.96 (2.76)	-0.86 (0.39)	0.0341
3	6.88 (4.13)	8.03 (4.37)	-1.15 (0.56)	0.0439	**4.99 (2.15)**	**6.21 (3.37)**	**-1.22 (0.38)**	**0.0023**
4	2.48 (1.08)	2.88 (1.01)	-0.40 (0.16)	0.0178	2.43 (0.96)	2.64 (0.86)	-0.21 (0.16)	0.1996
5	4.25 (1.82)	4.36 (1.76)	-0.11 (0.28)	0.6889	4.04 (1.60)	4.26 (1.49)	-0.22 (0.28)	0.4356
6	2.51 (1.36)	2.81 (1.40)	-0.30 (0.22)	0.1833	2.68 (1.36)	2.67 (1.21)	0.01 (0.25)	0.9657
7	3.22 (1.68)	3.44 (1.82)	-0.22 (0.29)	0.4610	3.47 (1.63)	3.16 (1.56)	0.32 (0.30)	0.3036
8	1.83 (1.12)	2.10 (1.04)	-0.27 (0.21)	0.1916	2.07 (0.97)	1.93 (0.95)	0.13 (0.18)	0.4738

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

Supplement for Question (2): "The Big Mac Index" is computed and reported every six months by *The Economist* http://www.economist.com/content/big-mac-index. It reports both the Raw Index and the Adjusted Index.

> The average price of a Big Mac in America in July 2015 was $4.79; in China it was only $2.74 at market exchange rates. So the "raw" Big Mac index says that the yuan was undervalued by 43% [(=100*(2.74 – 4.79)/4.79)].
>
> The adjusted index uses the "line of best fit" between Big Mac prices and GDP per person for 48 countries (plus the euro area). The difference between the price predicted by the [OLS] line for each country, given its income per person, and its actual price gives a supersized measure of currency under- and over-valuation.

The Economist posts the raw data that I downloaded (July 2015). The simple regression used to create the Adjusted Index uses GDP per capita in 2014 ($1,000s USD) to predict the price of a Big Mac hamburger in USD at market exchange rates. Here is the STATA summary of July 2015 data and the Big Mac price regression for a cross-section of 49 countries. ***Note:*** Some parts of the output are intentionally erased.

```
. summarize USD_price gdp_1000dollar;

      Variable |       Obs        Mean    Std. Dev.       Min        Max
---------------+--------------------------------------------------------
     USD_price |        49    3.588031    1.015294   1.832729   6.842105
gdp_1000dollar |        49    30.55946    22.97713     1.3427    97.0133


. regress USD_price gdp_1000dollar;

      Source |       SS       df       MS              Number of obs =      49
-------------+------------------------------           F(  1,    47) =   77.55
       Model |  30.8083359     1  30.8083359           Prob > F      =  0.0000
    Residual |  18.6711432    47  .397258365           R-squared     =
-------------+------------------------------           Adj R-squared =
       Total |  49.4794791    48  1.03082248           Root MSE      =  .63028

--------------------------------------------------------------------------------
     USD_price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
gdp_1000dollar |   .0348673   .0039593     8.81   0.000
         _cons |   2.522506   .1508211    16.73   0.000     2.219093    2.825919
--------------------------------------------------------------------------------
```

Remember, <u>only</u> your answers written on the test papers (in the designated space immediately after each question) will be graded. Any writing on this *Supplement* will *not* be graded.

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

Supplement for Question (3): Recall the readings and study materials assigned prior to this test for "Asiaphoria Meets Regression to the Mean," *NBER Working Paper 20573*, Oct. 2014, by Lant Pritchett and Larry Summers. All results in this *Supplement* use the more recent PWT 8.1 data.[4]

ABSTRACT Consensus forecasts for the global economy over the medium and long term predict the world's economic gravity will substantially shift towards Asia and especially towards the Asian Giants, China and India. While such forecasts may pan out, there are substantial reasons that China and India may grow much less rapidly than is currently anticipated. Most importantly, history teaches that abnormally rapid growth is rarely persistent, even though economic forecasts invariably extrapolate recent growth. Indeed, regression to the mean is the empirically most salient feature of economic growth. It is far more robust in the data than, say, the much-discussed middle-income trap. Furthermore, statistical analysis of growth reveals that in developing countries, episodes of rapid growth are frequently punctuated by discontinuous drop-offs in growth. Such discontinuities account for a large fraction of the variation in growth rates. We suggest that salient characteristics of China—high levels of state control and corruption along with high measures of authoritarian rule—make a discontinuous decline in growth even more likely than general experience would suggest. China's growth record in the past 35 years has been remarkable, and nothing in our analysis suggests that a sharp slowdown is inevitable. Still, our analysis suggests that forecasters and planners looking at China would do well to contemplate a much wider range of outcomes than are typically considered.

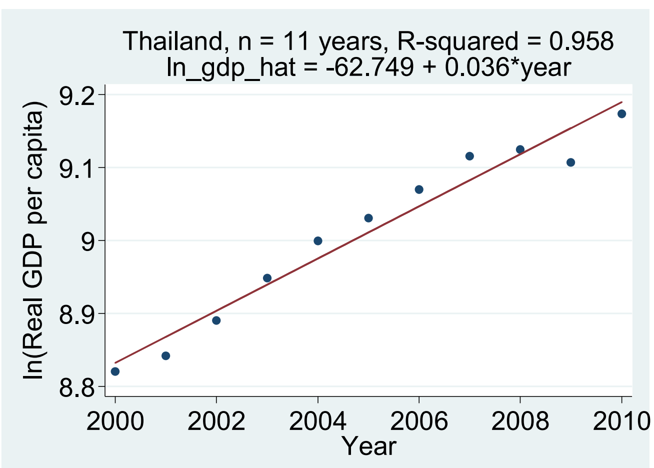


Table 1: Little persistence in cross-national growth rates across decades

Period 1	Period 2	Regression Coefficient	R-squared	N
PANEL A: *Adjacent decades*				
1950 – 60	1960 – 70	0.3375783	0.1236	66
1960 – 70	1970 – 80	0.4084345	0.1234	108
1970 – 80	1980 – 90	0.3225473	0.1138	142
1980 – 90	1990 – 00	0.2884994	0.1304	142
1990 – 00	2000 – 10	0.2051206	0.0562	142
PANEL B: *Two decades apart*				
1950 – 60	1980 – 90	-0.0475639	0.0020	66
1960 – 70	1990 – 00	0.1580633	0.0234	108
1970 – 80	2000 – 10	-0.0148128	0.0005	142
Source: Calculations based on PWT 8.1.				

[4] Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" forthcoming *American Economic Review*, available for download at www.ggdc.net/pwt. PWT 8.1 is an updated version of PWT 8.0, covering the same countries and period. Released on: April 13, 2015. (DOI: 10.15141/S5NP4S, Retrieved June 8, 2015.)

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

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

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

Multiplication rule: $P(A\ and\ B) = P(A|B)P(B) = P(B|A)P(A)$

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

Laws of expected value:

$E[c] = c$

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

$E[cX] = cE[X]$

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

Laws of variance:

$V[c] = 0$

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

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

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

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

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

Laws of covariance:

$COV[X,c] = 0$

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

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** $V[X] = \frac{(b-a)^2}{12}$

Sampling distribution of $\bar{X}$:

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

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

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

Sampling distribution of $\hat{P}$:

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

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

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

Sampling distribution of $(\hat{P}_2 - \hat{P}_1)$:

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

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

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

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

Sampling distribution of $(\bar{X}_1 - \bar{X}_2)$, independent samples:

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

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

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

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

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

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

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

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:** $\nu = n - 1$

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

SIMPLE REGRESSION:

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

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

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

Coefficient of determination: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$ $R^2 = (r)^2$ **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:** $s.e.(b_1) = s_{b_1} = \frac{s_e}{\sqrt{(n-1)s_x^2}}$

<u>Inference about the population slope</u>:

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

<u>Prediction interval for y at given value of x (x_g)</u>:

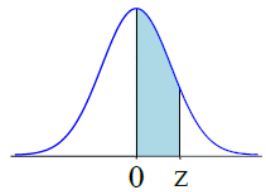
$\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}}$ **or** $\hat{y}_{x_g} \pm t_{\alpha/2} \sqrt{\left(s.e.(b_1)\right)^2 (x_g - \bar{X})^2 + \frac{s_e^2}{n} + s_e^2}$

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

<u>Confidence interval for predicted mean at given value of x (x_g)</u>:

$\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}}$ **or** $\hat{y}_{x_g} \pm t_{\alpha/2} \sqrt{\left(s.e.(b_1)\right)^2 (x_g - \bar{X})^2 + \frac{s_e^2}{n}}$ **Degrees of freedom:** $\nu = n - 2$

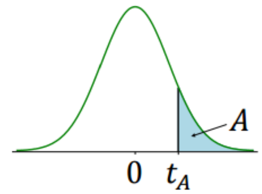
Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.



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

Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.



Critical Values of t:

ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
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	∞	1.282	1.645	1.960	2.326	2.576

Degrees of freedom: ν