ECO220Y, Test #2

June 26, 2015, 2:00 – 5:00

U of T e-mail: ____________________________________@mail.utoronto.ca

Surname (last name):	
Given name (first name):	
UTORID: (e.g. lihao8)	

Instructions:

- You have 3 hours. 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 5 written questions with varying point values worth a total of 110 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.
  - **You are encouraged to write in pencil and to use an eraser as needed.** This way you can make sure to fit your final answer (including work and reasoning) in the appropriate space.
- Clearly show your work. Make your reasoning clear.
- 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 asks 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 and Standard Normal 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) [40 pts] Read the entire *Supplement for Question (1)* and read the questions in Parts (a) – (f) below BEFORE answering any parts. Make sure to write your answers in the correct spots (i.e. do not jump ahead and answer questions not yet asked).

Page Pts:

(a) [3 pts] What is the main claim in "Asiaphoria Meets Regression to the Mean"? Answer with 1 – 2 sentences.

(b) [8 pts] What do the two graphs in the *Supplement for Question (1) (b)* show? Do the graphs show cross-sectional, time series, or panel data? What do the OLS results show? Fully interpret all numbers. Answer with 4 – 5 sentences. (*Note:* You are asked to describe/explain the results but *not* to discuss which conclusions to draw. That is asked later.)

(c) [4 pts] Compare and contrast the graphs and OLS results in the *Supplement for Question (1) (c)*, with the graphs and OLS results in the previous part. Answer with 2 – 3 sentences.

Page Pts:

(d) [9 pts] Use the graphs and OLS results in the *Supplement for Question (1) (d)* to support the authors' main warning. What do these three graphs show? How do they inform conclusions/forecasts from the graphs and OLS results in previous parts? Answer with 3 – 5 sentences.

(e) [8 pts] In the *Supplement for Question (1) (e) – (f)*, what do the results in PANEL A of the table mean? For *one* row of results, fully interpret all numbers. Conclusions? Answer with 3 – 5 sentences.

Page Pts:

(f) [8 pts] In the *Supplement for Question (1) (e) – (f)*, what do the results in PANEL B of the table mean? For *one* row of results, fully interpret all numbers. What do these results add beyond PANEL A? Answer with 3 – 5 sentences.

(2) [15 pts] The *Supplement for Question (2)* describes a population and a Monte Carlo simulation.

Page Pts:

(a) [5 pts] What is the probability that a randomly selected employee from that population makes more than $135,000? Answer with 1 – 2 sentences, showing your work/reasoning.

(b) [10 pts] If n = 30 were sufficiently large for the Central Limit Theorem, what would be the 99^{th} percentile? Why do the simulation results differ? How to interpret 139.8059? Answer with a quantitative analysis and 2 – 4 sentences.

(3) [18 pts] Review the *Supplement for Question (3)*. How big is the difference in callback rates for a Chinese name versus a Canadian-Chinese name? What is the margin of error? Fully interpret your findings. What can you conclude? Answer with a quantitative analysis and 4 – 5 sentences.

Page Pts:

(4) [19 pts] Review the *Supplement for Question (4)*, which revisits the sex ratios at birth research.

Page Pts:

(a) [4 pts] What is the hypothesis test to be repeated for each parity? Why does a one-tailed test makes sense? Answer with one set of formal hypotheses in standard notation and 1 sentence.

(b) [15 pts] For *each parity*, compute the P-value and assess the strength of the evidence in favor of the research hypothesis. Conclusions overall? Answer with a quantitative analysis, three P-values, and 3 – 4 sentences.

(5) [18 pts] Airlines often sell more tickets than there are seats (overbook). Usually some passengers do not show up (no-shows). If nearly all passengers show up, there will not be enough seats and some will be bumped (denied entry). The airline believes the rate of passenger no-shows is 0.06 (6 percent). It sold 900 tickets for a flight on an Airbus A380, which can hold 853 passengers. What is the chance that more than 853 passengers show up (i.e. some get bumped)? [Presume the independence assumption is reasonable.] Illustrate your answer with a fully-labelled graph where the horizontal axis is the *number of passengers that show up*. Answer with a quantitative analysis, a fully-labelled graph and 1 – 2 sentences.

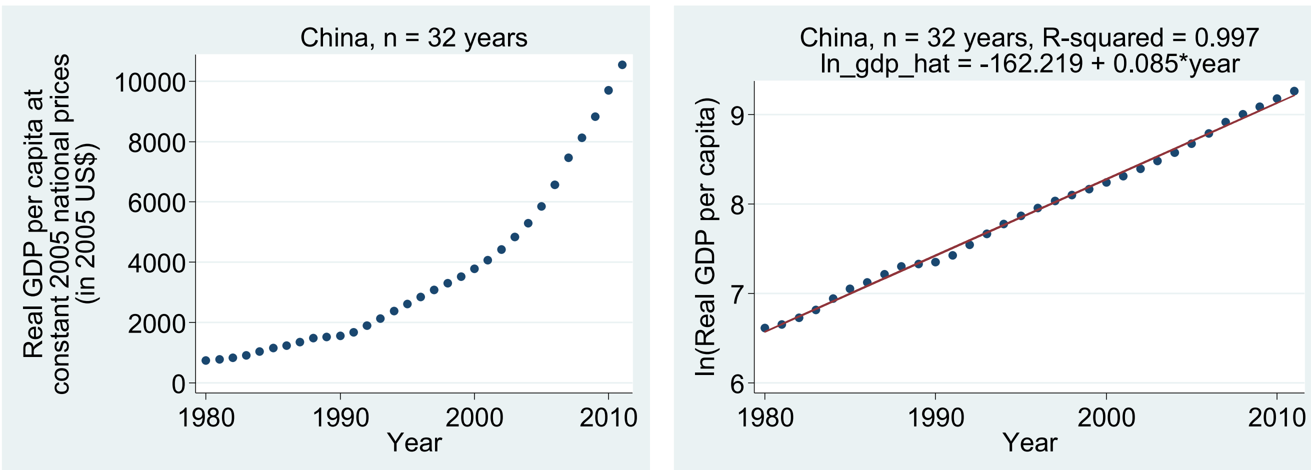
Page Pts:

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 and Standard Normal 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.

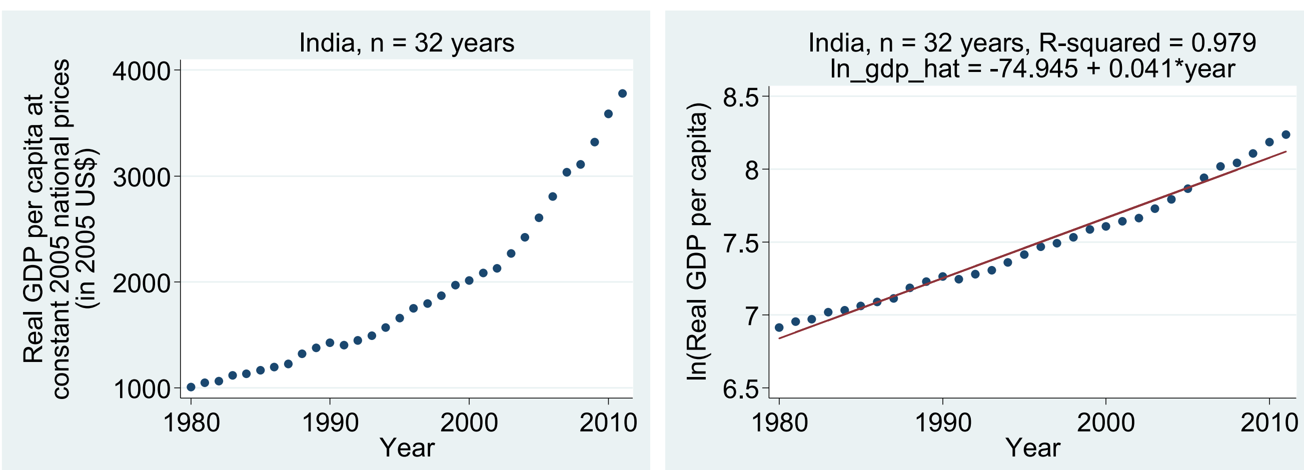
Supplement for Question (1): Recall the readings and study materials assigned prior to this test for "Asiaphoria Meets Regression to the Mean," *NBER Working Paper 20573*, Oct. 2014, by Lant Pritchett and Larry Summers. All results in this *Supplement* use the most recent PWT 8.1 data.[1]

Supplement for Question (1) (b):



OLS results: `ln(gdp)-hat = -162.219 + 0.085*year, R-squared = 0.997, n = 32`

Supplement for Question (1) (c):

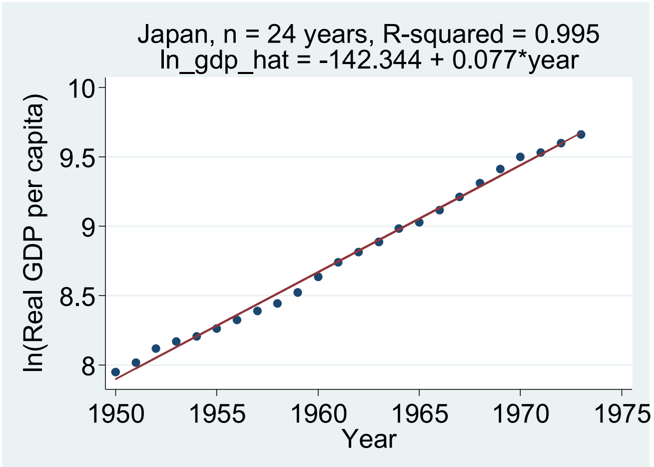


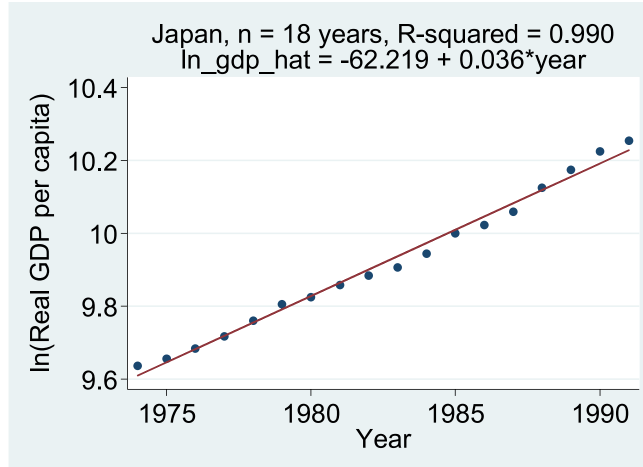
OLS results: `ln(gdp)-hat = -74.945 + 0.041*year, R-squared = 0.979, n = 32`

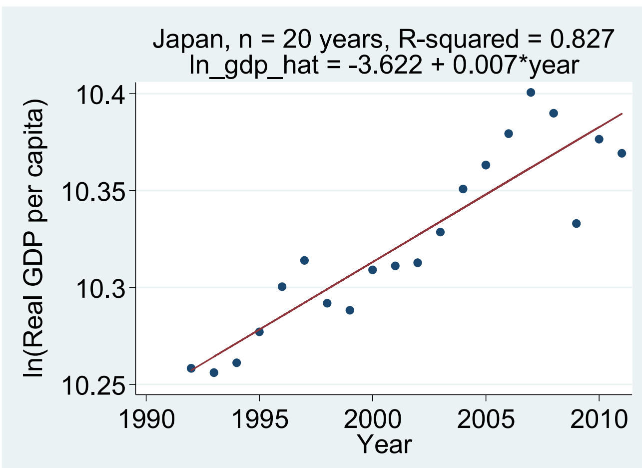
[1] Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" forthcoming *American Economic Review*, available for download at www.ggdc.net/pwt. PWT 8.1 is an updated version of PWT 8.0, covering the same countries and period. Released on: April 13,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

Supplement for Question (1) (d):







Note: Be sure to review the OLS results given in the title of each of these graphs.

Supplement for Question (1) (e) – (f):

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.				

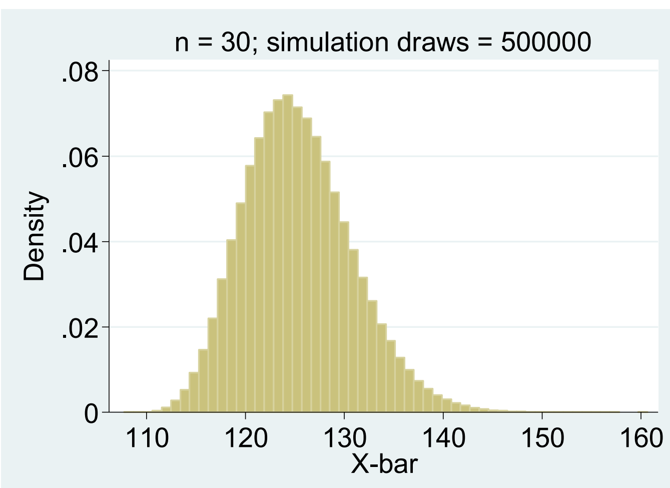
Supplement The pages of this supplement will *not* be graded: write your answers on the test papers.

Supplement for Question (2): Recall the publically available data for all ON public sector employees with salaries of $100,000 or more (http://www.fin.gov.on.ca/en/publications/salarydisclosure/pssd/). Consider the 98,942 employees in the 2014 disclosure of 2013 salaries that make $300,000 or less. A STATA summary shows the distribution of salaries (measured in $1,000s).

salary

	Percentiles	Smallest		
1%	100.2091	100		
5%	100.9725	100		
10%	102.0857	100	Obs	98942
25%	105.7196	100	Sum of Wgt.	98942
50%	115.1083		Mean	125.3419
		Largest	Std. Dev.	29.96436
75%	132.4765	299.9739		
90%	162.9707	300	Variance	897.863
95%	187.7392	300	Skewness	2.382879
99%	254.231	300	Kurtosis	10.12159

Supplement for Question (2) (b): Consider a Monte Carlo simulation. In each simulation draw, a random sample of 30 employees is drawn from the population of 98,942 employees. For each random sample, the sample mean is computed. 500,000 simulation draws are used. A histogram and STATA summary show the simulation results.



Sample mean

	Percentiles	Smallest		
1%	114.4613	107.7686		
5%	117.1032	108.0062		
10%	118.6248	108.2711	Obs	500000
25%	121.4469	108.2996	Sum of Wgt.	500000
50%	124.9363		Mean	125.3437
		Largest	Std. Dev.	5.470661
75%	128.795	156.2947		
90%	132.5799	156.7519	Variance	29.92813
95%	134.9801	157.5424	Skewness	.4398975
99%	139.8059	160.6329	Kurtosis	3.249203

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

Supplement for Question (3): In 2011, Philip Oreopoulos, a professor of economics at U of T, published a paper "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes" in the *American Economic Journal: Economic Policy*. Here is the paper's abstract:

> Thousands of randomly manipulated resumes were sent in response to online job postings in Toronto to investigate why immigrants, allowed in based on skill, struggle in the labor market. The study finds substantial discrimination across a variety of occupations against applicants with foreign experience or those with Indian, Pakistani, Chinese, and Greek names compared with English names. Listing language fluency, multinational firm experience, education from highly selective schools, or active extracurricular activities had no diminishing effect. Recruiters justify this behavior based on language skill concerns but fail to fully account for offsetting features when listed. (http://pubs.aeaweb.org/doi/pdfplus/10.1257/pol.3.4.148).

In this randomized field experiment, realistic but fake resumes were sent to real job postings. Consider resumes where the applicant had a Canadian education (e.g. a Rotman Commerce degree) and Canadian work experience. (This represents about half of all resumes sent as part of this experiment.) Each resume is *randomly assigned* a name. The author labels each name as either Canadian (e.g. "Emily Brown"), Canadian-Chinese, which is a Chinese last name and a Candian first name (e.g. "Eric Wang"), Chinese (e.g. "Min Liu"), Greek (e.g. "Lukas Minsopoulos"), Indian (e.g. "Shreya Sharma"), or Pakistani (e.g. "Ali Saeed"). The cross-tabulation below shows the ethnicity label of the name and callback records whether or not the potential employer ever called back the resume sender (no = 0 and yes = 1).

name_ethnicity	callback 0	callback 1	Total
Canadian	2,597	429	3,026
Canadian-Chinese	557	61	618
Chinese	890	108	998
Greek	329	37	366
Indian	1,047	120	1,167
Pakistani	429	37	466
Total	5,849	792	6,641

Supplement for Question (4): In "Sex ratios among Canadian liveborn infants of mothers from different countries" (*Canadian Medical Association Journal,* 2012, http://www.cmaj.ca/content/early/2012/04/16/cmaj.120165) the authors look at complete data for all Ontario births from 2002 – 2007. The authors analyze the data by where each mother was born and by parity (how many babies the woman has previously delivered). Consider babies born in Ontario to moms born in South Korea. For humans generally, the proportion of boys born is 0.512 (105 boys for every 100 girls).

Parity = 0 (first baby); South Korean mom; n = 1,815 babies; 934 male babies

Parity = 1 (second baby); South Korean mom; n = 1,439 babies; 786 male babies

Parity = 2 (third baby); South Korean mom; n = 344 babies; 175 male babies

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:** $r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n} z_{x_i} z_{y_i}}{n-1}$

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

SIMPLE REGRESSION: **OLS line:** $\hat{y}_i = b_0 + b_1 x_i$ $\quad b_1 = \frac{s_{xy}}{s_x^2} = r\frac{s_y}{s_x}$ $\quad b_0 = \bar{Y} - b_1\bar{X}$

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

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

$s_y^2 = \frac{SST}{n-1}$ **Coefficient of determination:** $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = (r)^2$

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: $P(A^C) = P(A') = 1 - P(A) \quad P(A^C|B) = P(A'|B) = 1 - P(A|B)$

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

Expected value: $E[X] = \mu = \sum_{all\,x} x p(x)$ **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:

$V[c] = 0$

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

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

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

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

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

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 $x = 0,1,2,\ldots,n$

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

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

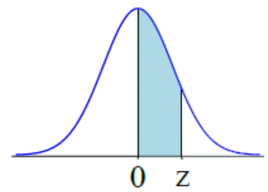
$\sigma_{\hat{P}_2 - \hat{P}_1}^2 = V[\hat{P}_2 - \hat{P}_1] = \frac{p_2(1-p_2)}{n_2} + \frac{p_1(1-p_1)}{n_1}$

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

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

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