Public Policy 529
Midterm Exam

Student ID number (8-digits): ____________________

1. In the city school district, the number of absences each day has a normal distribution with a mean of 150 and a standard deviation of 30.

   (a) What is the percentage of days that 190 or *fewer* students are absent?

   (b) What is the percentage of days that 135 students or *fewer* are absent?

   (c) For what percentage of days is the number of absent students in the range 105 to 195?

2. Explain the difference between a p-value and α and why we compare these two values in a significance test.

3. In the 2014 General Social Survey, respondents were asked about their attendance at a political rally. The distribution of the responses is given in the table below.

Attendance at political rally?	Freq.	Percent	Cumulative
Has attended in past year	100	8.1	8.1
Has attended in more distant past	269	21.6	29.7
Has not attended but might attend	479	38.5	68.2
Has not attended and never would	395	31.8	100.0
Total	1,243	100.0	

(a) Calculate the estimated proportion of respondents that have attended a political rally and the 95% confidence interval around this estimate.

(b) Interpret the confidence interval.

4. True or False? When the population distribution of y is highly skewed, it is always inappropriate to use the t distribution to represent the sampling distribution of $\bar{y}$. Explain.

5. The Health Monitoring Reform Survey asked respondents whether they had health insurance coverage of some kind (yes/no) and about their employment status. The resulting data are presented in the table below. The cells contain frequencies.

Insured?	Employee	Self-Employed	Retired	Not Working	Total
No	324	97	23	280	724
Yes	4,286	464	547	1,621	6,918
Total	4,610	561	570	1,901	7,642

(a) What are the measurement levels of these two variables?

(b) What is P(Not Working and Not Insured)?

(c) What is P(Insured | Self-Employed)?

(d) Are health insurance coverage and work status independent of each other? Demonstrate mathematically.

6. You are asked to evaluate the results of a program to increase the reading test scores of elementary school-aged children over the summer. Among the 780 children who took part in the program, the mean reading test score at the end of summer was 240 with a standard deviation of 70. Perform a test of statistical significance (α=.05) of the assumption that the "true" mean is 230. Report your test statistic, the critical value of the test statistic, and p-value.

7. Two researchers measure crime rates in the city for various types of crime during the past year. The first researcher obtains police department data for each type of crime reported during this time period. The second researcher conducts a survey of 120 randomly-selected city residents to identify the proportions that were victims of each type of crime during the past year. Evaluating each strategy separately, identify possible sources of measurement error and explain how each source of error would affect the reliability and/or validity of the measurement strategy.

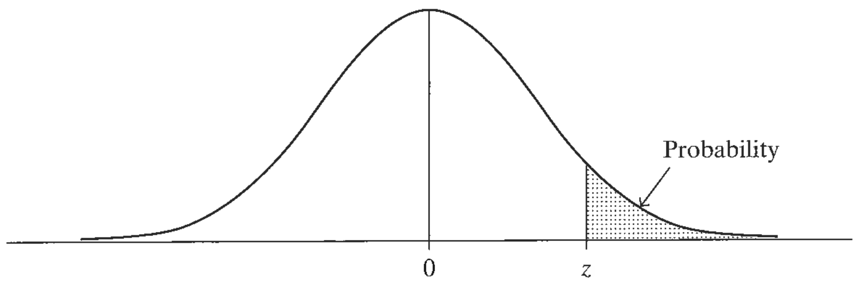
8. Suppose that, in the population, the mean years of education is 14 with a standard deviation of 2.3.

   (a) A researcher, who doesn't know the population mean, is going to take a random sample of size 1,200. In what range is the sample mean expected to be in 90% of the time?

   (b) Continuing with the above scenario, suppose the mean of the sample is 14.2 with a standard deviation of 2. Construct a 90% confidence interval for the estimate and interpret it.

Space for Work

TABLE A: Normal curve tail probabilities. Standard normal probability in right-hand tail (for negative values of z, probabilities are found by symmetry)



	Second Decimal Place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000233									
4.0	.0000317									
4.5	.00000340									
5.0	.000000287									

Source: R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).

List of Formulas

Descriptive and Distributional Statistics

$$\bar{y} = \frac{\sum y_i}{n}$$

$$s^2 = \frac{\sum(y_i - \bar{y})^2}{n-1}$$

$$Z = \frac{y - \mu_y}{\sigma}$$

$$IQR = Q_3 - Q_1$$

$$SS = \sum(y_i - \bar{y})^2$$

$$s = \sqrt{\frac{\sum(y_i - \bar{y})^2}{n-1}}$$

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}$$

Probability

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

$$P(A \text{ and } B) = P(A) \times P(B|A)$$

$$P(\sim A) = 1 - P(A)$$

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$P(A \text{ and } B) = P(A) \times P(B)$$

Confidence Intervals and Significance Tests

$$Z \text{ or } t = \frac{\bar{y} - \mu_0}{\hat{\sigma}_{\bar{y}}}$$

$$\hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

$$Z = \frac{\hat{\pi} - \pi_0}{\hat{\sigma}_{\pi_0}}$$

$$\text{c.i.} = \bar{y} \pm t \cdot \hat{\sigma}_{\bar{y}}$$

$$\hat{\sigma}_{\bar{y}} = \frac{s}{\sqrt{n}}$$

$$\text{c.i.} = \bar{y} \pm Z \cdot \hat{\sigma}_{\bar{y}}$$

$$\hat{\sigma}_{\pi_0} = \sqrt{\frac{\pi_0(1-\pi_0)}{n}}$$

$$\text{c.i.} = \hat{\pi} \pm Z \cdot \hat{\sigma}_{\hat{\pi}}$$