PSCI 2300 • Intro to Poli Sci Research Methods (Fall 2019)
Dr. Paul Hensel
Review Sheet for Exam #2

Exam Format

• This examination will feature 30 multiple choice questions (worth 1 point each) and 5 short answer questions (worth 2 points each). No calculators will be needed; the short answer questions will require you to define relevant terms, interpret results that are provided for you, or apply statistical tables (which will be provided for you) to answer questions, and won't require advanced math like square roots.
• All needed exam materials (such as Scantron sheets) will be provided for you, except for your preferred pencil (and eraser, if needed).
• No books, notes, or other materials will be allowed during the exam; copies of the relevant formulas and tables will be included with the exam.
• Note that no bathroom breaks are permitted during the exam -- once you leave the room, you can't return without a five letter grade penalty.
• Also, be on time -- once the first person finishes and leaves the room, anybody else who enters will suffer an automatic deduction of five letter grades; be aware that students sometimes finish exams like this in as little as 5-10 minutes.

Advice on Studying

• The number of questions on each topic in the exam will (approximately) reflect the relative time spent on each topic in the course. A topic that was covered over two lectures should thus have approximately (but perhaps not exactly) twice as many exam questions as a topic that only lasted for one lecture.
• The exam will draw from both the assigned readings (including both textbook and assigned journal articles/book chapters) and the lecture notes. As many as 1/3 of the exam questions will be drawn from materials in the readings that were not covered at all in lecture, with the rest reflecting topics that were only covered in lecture or were covered in both lecture and the readings. If you missed one or more days of class, be sure to get a copy of those notes from somebody who was there.
--Note that in the past, questions drawn from the readings (even on concepts or topics specifically listed on the review sheet) have produced the lowest scores of any questions on the test, so you should take special care to study these topics on the review sheet.
• The exam will not be written with the intention of fooling students with trick questions or with the goal of failing as many students as possible. The main goal of this course is to provide students with an understanding of how Political Science research works, so the exam questions will reflect this goal.
• The list of topics on this review sheet is not legally binding; these are just suggestions for the most important topics that are most likely to be on the test (some of which may not actually appear on the test). If you understand all of these topics you are much more likely to do well on the test.

Examples of Possible Short Answer Questions

• Identification/calculation of mode, median, and mean
• Interpretation of descriptive statistics (if a statistic or graph is provided for you, you will need to interpret it)
• Definition and interpretation of **α** (alpha)
• Usage and interpretation of Z-scores using table from book
• Definition and interpretation of confidence intervals
• Null hypotheses (if a research hypothesis is provided for you, you will need to write the appropriate null hypothesis)
• Definition and interpretation of statistical significance
• Types of hypotheses (1-tailed vs. 2-tailed, directional vs. non-directional): what are the differences between them, and how do these differences affect statistical significance?
• Usage and interpretation of t-statistics using table from book
• Calculation and interpretation of p-values using table from book (for any method that we studied)

Topics Covered in This Portion of the Course

Descriptive Statistics

• Nominal, Ordinal, Interval, and Ratio-level variables
• Additive indexes
• Descriptive vs. inferential stats
• Type of descriptive stats: Frequency, central tendency, dispersion
• Descriptive stats for nominal-level variables: meaning, interpretation, when to use
    --Frequency distribution
    --Proportion and percentage: meaning and interpretation
    --Bar graph/chart: meaning and interpretation
    --Pie graph/chart: meaning and interpretation
    --Mode: meaning and interpretation
    --Dispersion (eyeball/inter-ocular trauma): meaning and interpretation
• Descriptive stats for ordinal-level variables: meaning, interpretation, when to use
    --(relevant stats discussed under nominal-level variables)
    --Cumulative frequency and percentile: meaning and interpretation
    --Median: meaning and interpretation
• Descriptive stats for interval-/ratio-level variables: meaning, interpretation, when to use
    --(relevant stats discussed under nominal-/ordinal-level variables)
    --Histogram: meaning and interpretation
    --Line graph: meaning and interpretation
    --Mean: meaning and interpretation (including impact of skewness, possible solutions)
    --Standard deviation: meaning and interpretation

Additional Topics from the Readings (Pollock chapter 2 "Describing Variables" and Summary sections, Pollock chapter 6 "Variation Revisited: The Standard Deviation" section)

• More detail on additive indexes and Likert scales
• More detail on meaning of central tendency and dispersion
• More detail on interquartile range
• Box plots

Sampling and Inferential Statistics

• Basic idea of survey research: obtain information from a small representative sample (sample statistics), to make inferences about a larger population (population parameters)

• Approaches to representative sample: random samples, stratified/quota samples, weighting, NOT convenience sampling
• Approaches to collecting survey information: face to face, telephone, mail-in, online
• Survey advantages, internal and external validity
• Survey disadvantages/weaknesses: importance of question wording, assume honest answers, lack of representative samples, low response rate, difficulty in weighting results
• Populations (and pop. parameters) vs. samples (and sample statistics), and purpose of sampling
• Goal of random sampling
• Systematic errors in sampling: selection bias, response bias, poorly written Qs, social desirability
• Random sampling error: meaning, intuition (variation component / sample size component)
• Normal distribution/normal curve: shape, mean=mode=median, symmetric/no skew
• Probability density function and normal distribution (% cases within 1-2-3 SDs)
• Z-score: meaning, calculation
• Use of standard normal distribution with Z-scores (Table 6-3 in book)
• Central Limit Theorem and characteristics of sampling distribution
• Pop. parameters: mean=μ, standard deviation=σ
• Sample statistics: mean=$\bar{x}$, standard deviation=s
• Meaning of α (alpha)
• 95% confidence interval for mean: meaning, calculation (sample mean +/- 2 standard units)
• Student's t vs. Z: when each should be used
• Use of t distribution (Table 6-4 in book)
• 95% confidence interval for proportion: meaning, calculation

Additional Topics from the Readings (Pollock chapter 6, Cohn article about 2016 polls)
• More detail on random sampling (sampling frames, selection bias, response bias, random selection)
• Inference about sample proportions

Hypothesis Testing
• Research hypotheses, null hypotheses, and testing
• Directional vs. non-directional hypotheses (and corresponding null hypotheses)
• Statistical significance, p-values: definition and role in hypothesis testing
• 1-tailed vs 2-tailed tests: when used, how they impact significance testing
• Type 1 and Type II errors: definition, when each is more/less likely
• Limitations of statistical significance: covariation vs. causation, inflation as N increases, difference from substantive significance, role of p-hacking
• Difference of means tests: what is actually being tested (research & null hypotheses)
• Confidence intervals and simple difference of means tests
• General procedure for more precise difference of means tests: assumptions, hypotheses, critical values, computing t-statistic, making the decision
• Critical values: definition, role in hypothesis testing
• ANOVA: what is actually being tested (research & null hypotheses)
• General procedure for ANOVA: assumptions, hypotheses, critical values, computing F-statistic, making the decision

Additional Topics from the Readings (Pollock chapter 7 Intro, "Statistical Significance", and "Comparing Two Sample Means" sections)
• Error bar chart/graph
• Eyeball test of statistical significance
• Compairing sample proportions

Formula Sheet for Exam #2
[this formula sheet and these statistical tables will be included with the exam]

Descriptive Statistics

• Proportion = **f / N** *(where f=frequency of cases in this category, and N=total sample size)*

• Percentage (%) = **(f/N)*100**

• Mean: $\bar{x} = \frac{\Sigma x_i}{N}$ *(where x_i = the individual observation, N=sample size)*

• Calculating standard deviation:

(1) Deviation = $x_i - \bar{x}$

(2) Squared deviation = $(x_i - \bar{x})^2$

(3) Total Sum of Squares/TSS (Sum of Squared Deviations/SSD) = $\Sigma (x_i - \bar{x})^2$

(4) Variance: $\sigma^2 = \frac{TSS}{n} = \frac{\Sigma (x_i - \bar{x})^2}{n}$

(5) Standard deviation: $\sigma = \sqrt{\sigma^2}$

Sampling and Inference

• Z score: $Z = \frac{\text{Deviation from mean}}{\text{Sample SD}} = \frac{x_i - \bar{x}}{s}$

(where x_i=score for this observation, $\bar{x}$= mean for sample, s= std dev for sample)

• Std error of sample mean = random sampling error = $\frac{\text{Variation component}}{\text{Sample size component}} = \frac{s}{\sqrt{n}}$

(where s=std dev of sample, n=sample size)

• 95% Confidence interval for mean = $\bar{x}$ **+/- 2 S.E.** *(where $\bar{x}$=sample mean, S.E.= $\frac{s}{\sqrt{n}}$)*

Hypothesis Tests for Interval/Ratio Variables

• Difference of means: $t_{Observed} = \frac{\text{Difference between sample means}}{\text{Dispersion of sample}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{(s_{xbar1}^2 + s_{xbar2}^2)}}$

(where $\bar{x}_1$ and $\bar{x}_2$=sample means; s_{xbar1} and s_{xbar2}=sample std errors: $s_{xbar} = \frac{2}{\sqrt{n}}$)

• ANOVA: $F = \frac{\text{Avg variation between}}{\text{Avg variation within}} = \frac{\text{Explained variance (IV)}}{\text{Unexplained variance (error)}}$

Table 6-3 Proportions of the Normal Curve above the Absolute Value of Z

First digit and first decimal of Z	Second decimal of 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	.0721	.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	.0351	.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	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

Table 6-4 The Student's *t*-Distribution

	Area under the curve			
Degrees of freedom	*.10*	*.05*	*.025*	*.01*
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528
21	1.323	1.721	2.080	2.518
22	1.321	1.717	2.074	2.508
23	1.319	1.714	2.069	2.500
24	1.318	1.711	2.064	2.492
25	1.316	1.708	2.060	2.485
26	1.315	1.706	2.056	2.479
27	1.314	1.703	2.052	2.473
28	1.313	1.701	2.048	2.467
29	1.311	1.699	2.045	2.462
30	1.310	1.697	2.042	2.457
40	1.303	1.684	2.021	2.423
60	1.296	1.671	2.000	2.390
90	1.291	1.662	1.987	2.368
100	1.290	1.660	1.984	2.364
120	1.289	1.658	1.980	2.358
1,000	1.282	1.646	1.962	2.330
Normal (*Z*)	1.282	1.645	1.960	2.326