

Week 20

AP Statistics

Homework bundle for this week

本周作业合集

exercises.lol

Contents 目录

Topic 6 quiz	3
Exercise sheet 6.1	6
Exercise sheet 6.2	9
Past-paper questions 6.2	13
Exercise sheet 6.3	23
Past-paper questions 6.3	26
Exercise sheet 6.4	35
Past-paper questions 6.4	39
Exercise sheet 6.5	42
Exercise sheet 6.6	45
Past-paper questions 6.6	48
Exercise sheet 6.7	51

Past-paper questions 6.7	54
Exercise sheet 6.8	62
Past-paper questions 6.8	65
Exercise sheet 6.9	66
Exercise sheet 6.10	69
Past-paper questions 6.10	72
Exercise sheet 6.11	75
Past-paper questions 6.11	79

6. Inference for Categorical Data: Proportions

Starter Quiz · 课前小测

AP Statistics

Name: _____ Score: _____

1. Inference about a proportion mainly does what?

- ☐ quantifies the uncertainty in a sample statistic
- ☐ removes all uncertainty
- ☐ guarantees the true p
- ☐ makes the sample bigger

2. $n = 100$, $\hat{p} = 0.60$. Compute the standard error $\sqrt{(\hat{p}(1-\hat{p})/n)}$. Two decimals.

3. The correct interpretation of a 95% confidence level is...

- ☐ about 95% of intervals built this way capture p over many samples
- ☐ there's a 95% probability p is in this one interval
- ☐ 95% of the data lie in the interval
- ☐ the sample is 95% accurate

4. The null hypothesis for a one-proportion test has the form...

- ☐ $H_0: p = p_0$
- ☐ $H_a: p \neq p_0$
- ☐ $H_0: \hat{p} = 0$
- ☐ $H_a: p > 1$

5. The p-value is the probability of a result at least as extreme as observed, assuming...

- ☐ H_0 is true
- ☐ H_a is true
- ☐ the sample is biased
- ☐ $\hat{p} = 0.5$

6. The decision rule for a significance test is: reject H_0 when...

- ☐ $p\text{-value} \leq \alpha$
- ☐ $p\text{-value} > \alpha$
- ☐ $z = 0$
- ☐ $\hat{p} = p_0$

7. $\hat{p}_1 = 0.60$, $\hat{p}_2 = 0.50$. Find the point estimate $\hat{p}_1 - \hat{p}_2$.

8. $\hat{p}_1 = 0.30$, $\hat{p}_2 = 0.20$, pooled $SE \approx 0.061$. Compute $z = (\hat{p}_1 - \hat{p}_2)/SE$. Two decimals.

9. The large counts condition ($n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$) justifies using a normal model.

- ☐ True ☐ False

10. With $SE \approx 0.049$ and $z^* = 1.96$, find the margin of error $z^* \cdot SE$. Two decimals.

AP Statistics

1. **quantifies the uncertainty in a sample statistic**

Inference measures uncertainty honestly — it doesn't remove it.

2. **0.05**

$$\sqrt{(0.6 \cdot 0.4/100)} = \sqrt{0.0024} \approx 0.049 \approx 0.05.$$

3. **about 95% of intervals built this way capture p over many samples**

Confidence describes the method over repeated sampling.

4. **H0: p = p0**

H0 states a specific value p0 (the status quo).

5. **H0 is true**

p-value is computed under the assumption that H0 is true.

6. **p-value $\leq \alpha$**

Reject when the p-value is at most the significance level.

7. **0.1**

$$0.60 - 0.50 = 0.10.$$

8. **1.64**

$$0.10 / 0.061 \approx 1.64.$$

9. **True**

Large counts make the sampling distribution roughly normal.

10. **0.1**

$$1.96 \times 0.049 \approx 0.096 \approx 0.10.$$

6.1 Introducing Statistics: Why Be Normal?

Name: _____ Class: _____ Date: _____ **Total: 8 marks**

KEY WORDS inference 推断 • normal 正态 • large counts condition 大计数条件
• confidence interval 置信区间

Objective

Build the skills to answer exam questions on **why the normal model matters for inference**.

You must be able to:

- explain that the **normal** 正态 approximation underlies inference for proportions
- state the **large counts condition** 大计数条件
- describe what inference estimates from a sample

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ Why normal?

Because the sampling distribution of $\hat{p}$ is approximately **normal** (when conditions are met), we can use z-values to build confidence intervals and run tests about a population proportion.

■ The large counts condition

The normal model applies when $np \geq 10$ **and** $n(1 - p) \geq 10$. Inference then uses the sample to estimate the unknown population proportion p .

2 Practice

2.1 State why the normal distribution is central to inference for proportions. [1]

2.2 State the large counts condition for a proportion. [1]

2.3 State what inference uses a sample to estimate. [1]

3 Exam-style questions

3.1 Inference for proportions relies on the _____ distribution. [1]

- **A** binomial
 - **B** normal
 - **C** uniform
 - **D** skewed
-

3.2 The large counts condition requires np and $n(1 - p)$ to each be at least [1]

- **A** 1
 - **B** 5
 - **C** 10
 - **D** 30
-

3.3 A sample proportion is used to estimate a population proportion.

(a) Name the distribution used for inference. [1]

(b) State the condition for it to apply. [1]

(c) State what we are estimating. [1]

Solutions

2.1 the sampling distribution of $\hat{p}$ is approximately normal, allowing z -based inference.

2.2 $np \geq 10$ and $n(1 - p) \geq 10$.

2.3 the unknown population proportion p .

3.1 **B**.

3.2 **C**.

3.3 (a) the normal distribution. (b) $np \geq 10$ and $n(1 - p) \geq 10$. (c) the population proportion p .

6.2 Constructing a Confidence Interval for a Population Proportion

Name: _____ Class: _____ Date: _____ Total: 9 marks

KEY WORDS confidence interval 置信区间 • margin of error 误差幅度

Objective

Build the skills to answer exam questions on a **confidence interval for a population proportion**.

You must be able to:

- construct a confidence interval $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
- use the appropriate **critical value** z^*
- interpret the interval

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ The confidence interval

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

where the **margin of error** is z^* times the standard error. For 95% confidence, $z^* = 1.96$.

■ Interpretation

"We are 95% confident that the true proportion p lies within the interval."

■ Example

$$\hat{p} = 0.6, n = 100, 95\%: \text{ME} = 1.96 \sqrt{\frac{0.6(0.4)}{100}} = 1.96(0.049) = 0.096.$$

2 Practice

2.1 Write the formula for a confidence interval for a proportion. [1]

2.2 State the z^* value for a 95% confidence interval. [1]

2.3 A poll gives $\hat{p} = 0.5$ with $n = 100$. Find the 95% margin of error. [2]

3 Exam-style questions

3.1 A confidence interval for p has the form [1]

- A $\hat{p} \pm \text{ME}$
- B $\hat{p} \times \text{ME}$
- C $\hat{p}/\text{ME}$
- D ME only

3.2 For a 95% confidence interval, z^* is about [1]

- A 1.28
- B 1.645
- C 1.96
- D 2.58

3.3 $\hat{p} = 0.4$, $n = 100$, 95% confidence.

(a) Find the margin of error.

[2]

(b) Give the confidence interval.

[1]

Solutions

2.1 $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

2.2 1.96.

2.3 $ME = 1.96 \sqrt{\frac{0.5(0.5)}{100}} = 1.96(0.05) = 0.098$.

3.1 A.

3.2 C.

3.3 (a) $ME = 1.96 \sqrt{\frac{0.4(0.6)}{100}} = 1.96(0.049) = 0.096$. (b) (0.304, 0.496).

Name:

AP Statistics: 2022

A survey conducted by a national research center asked a random sample of 920 teenagers in the United States how often they use a video streaming service. From the sample, 59% answered that they use a video streaming service every day.

- (a) Construct and interpret a 95% confidence interval for the proportion of all teenagers in the United States who would respond that they use a video streaming service every day.

GO ON TO THE NEXT PAGE.

Name:

AP Statistics: 2022

Continue your response to **QUESTION 4** on this page.

- (b) Based on the confidence interval in part (a), do the sample data provide convincing statistical evidence that the proportion of all teenagers in the United States who would respond that they use a video streaming service every day is not 0.5? Justify your answer.

GO ON TO THE NEXT PAGE.

**STATISTICS
SECTION II**

Part B

Question 6

Spend about 25 minutes on this part of the exam.

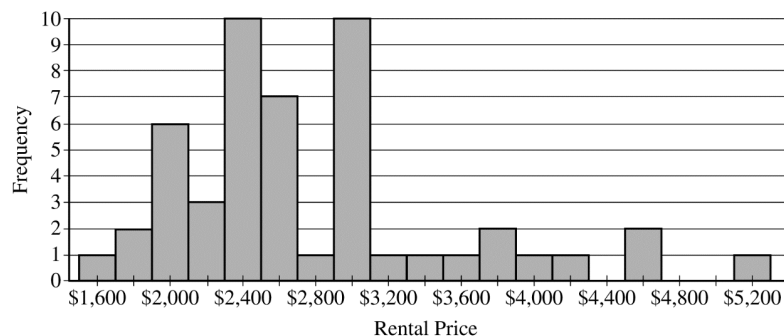
Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Emma is moving to a large city and is investigating typical monthly rental prices of available one-bedroom apartments. She obtained a random sample of rental prices for 50 one-bedroom apartments taken from a Web site where people voluntarily list available apartments.

(a) Describe the population for which it is appropriate for Emma to generalize the results from her sample.

The distribution of the 50 rental prices of the available apartments is shown in the following histogram.



- (b) Emma wants to estimate the typical rental price of a one-bedroom apartment in the city. Based on the distribution shown, what is a disadvantage of using the mean rather than the median as an estimate of the typical rental price?
- (c) Instead of using the sample median as the point estimate for the population median, Emma wants to use an interval estimate. However, computing an interval estimate requires knowing the sampling distribution of the sample median for samples of size 50. Emma has one point, her sample median, in that sampling distribution. Using information about rental prices that are available on the Web site, describe how someone could develop a theoretical sampling distribution of the sample median for samples of size 50.

Because Emma does not have the resources to develop the theoretical sampling distribution, she estimates the sampling distribution of the sample median using a process called bootstrapping. In the bootstrapping process, a computer program performs the following steps.

- Take a random sample, with replacement, of size 50 from the original sample.
- Calculate and record the median of the sample.
- Repeat the process to obtain a total of 15,000 medians.

Emma ran the bootstrap process, and the following frequency table is the bootstrap distribution showing her results of generating 15,000 medians.

Bootstrap Distribution of Medians					
Median	Frequency	Median	Frequency	Median	Frequency
2,345	1	2,585	1	2,825	247
2,390	13	2,587.5	171	2,837.5	7
2,395	18	2,600	22	2,847.5	1
2,400	56	2,612.5	1,190	2,872.5	317
2,445	4	2,625	174	2,885	10
2,447.5	56	2,672.5	5	2,950	700
2,450	55	2,675	1,924	2,962.5	93
2,475	3	2,687.5	1,341	2,972.5	6
2,495	66	2,700	2,825	2,975	65
2,497.5	136	2,735	35	2,985	12
2,500	1,899	2,747.5	619	2,987.5	1
2,522.5	2	2,750	2	2,995	6
2,525	945	2,795	278	3,000	2
2,550	1,673	2,812.5	16	3,062.5	3

The bootstrap distribution provides an approximation of the sampling distribution of the sample median. A confidence interval for the median can be constructed using a percentage of the values in the middle of the bootstrap distribution.

- (d) Use the frequency table to find the following.
- Value of the 5th percentile:
 - Value of the 95th percentile:
- (e) Find the percentage of bootstrap medians in the table that are equal to or between the values found in part (d).
- (f) Use your values from parts (d) and (e) to construct and interpret a confidence interval for the median rental price.

STOP

END OF EXAM

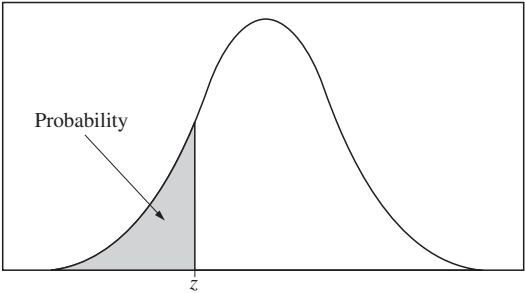


Table entry for z is the probability lying below z .

Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

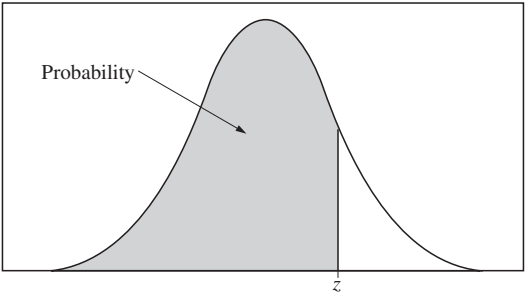


Table entry for z is the probability lying below z .

Table A (Continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^* .

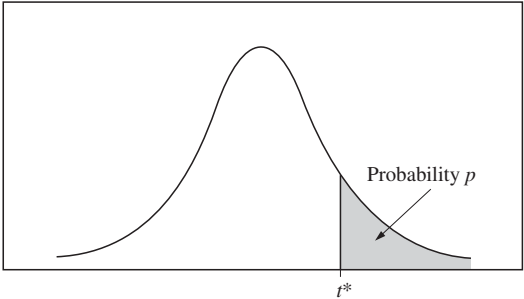


Table B t distribution critical values

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
50% 60% 70% 80% 90% 95% 96% 98% 99% 99.5% 99.8% 99.9%												
Confidence level C												



Table entry for p is the point (χ^2) with probability p lying above it.

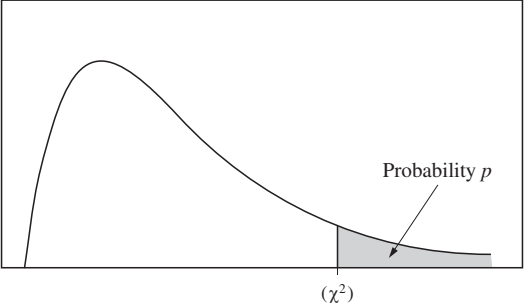


Table C χ^2 critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2



2. An environmental science teacher at a high school with a large population of students wanted to estimate the proportion of students at the school who regularly recycle plastic bottles. The teacher selected a random sample of students at the school to survey. Each selected student went into the teacher's office, one at a time, and was asked to respond yes or no to the following question.

Do you regularly recycle plastic bottles?

Based on the responses, a 95 percent confidence interval for the proportion of all students at the school who would respond yes to the question was calculated as (0.584, 0.816).

- (a) How many students were in the sample selected by the environmental science teacher?
- (b) Given the method used by the environmental science teacher to collect the responses, explain how bias might have been introduced and describe how the bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question.
- (c) The statistics teacher at the high school was concerned about the potential bias in the survey. To obtain a potentially less biased estimate of the proportion, the statistics teacher used an alternate method for collecting student responses. A random sample of 300 students was selected, and each student was given the following instructions on how to respond to the question.
 - In private, flip a fair coin.
 - If heads, you must respond no, regardless of whether you regularly recycle.
 - If tails, please truthfully respond yes or no.
 - (i) What is the expected number of students from the sample of 300 who would be required to respond no because the coin flip resulted in heads?
 - (ii) The results of the sample showed that 213 of the 300 selected students responded no. Based on the results of the sample, give a point estimate for the proportion of all students at the high school who would respond yes to the question.

2. The manager of a local fast-food restaurant is concerned about customers who ask for a water cup when placing an order but fill the cup with a soft drink from the beverage fountain instead of filling the cup with water. The manager selected a random sample of 80 customers who asked for a water cup when placing an order and found that 23 of those customers filled the cup with a soft drink from the beverage fountain.
- (a) Construct and interpret a 95 percent confidence interval for the proportion of all customers who, having asked for a water cup when placing an order, will fill the cup with a soft drink from the beverage fountain.
 - (b) The manager estimates that each customer who asks for a water cup but fills it with a soft drink costs the restaurant \$0.25. Suppose that in the month of June 3,000 customers ask for a water cup when placing an order. Use the confidence interval constructed in part (a) to give an interval estimate for the cost to the restaurant for the month of June from the customers who ask for a water cup but fill the cup with a soft drink.

5. A polling agency showed the following two statements to a random sample of 1,048 adults in the United States.

Environment statement: Protection of the environment should be given priority over economic growth.

Economy statement: Economic growth should be given priority over protection of the environment.

The order in which the statements were shown was randomly selected for each person in the sample. After reading the statements, each person was asked to choose the statement that was most consistent with his or her opinion. The results are shown in the table.

	Environment Statement	Economy Statement	No Preference
Percent of sample	58%	37%	5%

- (a) Assume the conditions for inference have been met. Construct and interpret a 95 percent confidence interval for the proportion of all adults in the United States who would have chosen the economy statement.
- (b) One of the conditions for inference that was met is that the number who chose the economy statement and the number who did not choose the economy statement are both greater than 10. Explain why it is necessary to satisfy that condition.
- (c) A suggestion was made to use a two-sample z -interval for a difference between proportions to investigate whether the difference in proportions between adults in the United States who would have chosen the environment statement and adults in the United States who would have chosen the economy statement is statistically significant. Is the two-sample z -interval for a difference between proportions an appropriate procedure to investigate the difference? Justify your answer.

2. To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

The owner selected a random sample of bills and found that only 15 percent of them received discounts. A confidence interval for p , the proportion of bills that will receive a discount in the long run, is 0.15 ± 0.06 . All conditions for inference were met.

- (a) Consider the confidence interval 0.15 ± 0.06 .
 - (i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.
 - (ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2? Justify your answer.

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

- (b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for p with the same confidence level as that of the original interval.
- (c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

6.3 Justifying a Claim Based on a Confidence Interval for a Population Proportion

Name: _____ Class: _____ Date: _____ Total: 8 marks

KEY WORDS confidence interval 置信区间 • confidence level 置信水平

Objective

Build the skills to answer exam questions on **justifying a claim from a confidence interval for a proportion**.

You must be able to:

- use a confidence interval to judge whether a claimed value is **plausible**
- recognise that a value **inside** the interval is plausible and one **outside** is not supported
- state the meaning of the confidence level

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ Using a confidence interval to judge a claim

A confidence interval gives the plausible values of p :

- if a claimed value is **inside** the interval, the claim is **plausible**;
- if it is **outside**, the data do **not support** that claim.

■ Example

A 95% interval of (0.52, 0.58) does **not** contain 0.5, so " $p = 0.5$ " is not supported, but "a majority ($p > 0.5$)" is plausible.

2 Practice

2.1 State how a confidence interval can be used to judge a claim. [1]

2.2 A 95% CI for p is (0.45, 0.55). State whether the claim " $p = 0.5$ " is plausible. [1]

2.3 State what it means if a claimed value lies outside the interval. [1]

3 Exam-style questions

3.1 A claimed value that lies inside a confidence interval is [1]

- A rejected
- B plausible
- C impossible
- D certainly correct

3.2 A 95% CI for p is (0.60, 0.70). A claim that $p = 0.50$ is [1]

- A plausible
- B not supported (it is outside the interval)
- C certainly true
- D unknown

3.3 A 95% CI for the proportion supporting a policy is (0.52, 0.58).

(a) State whether "a majority ($p > 0.5$) support it" is plausible. [1]

(b) State whether $p = 0.5$ is inside the interval. [1]

(c) State the confidence level. [1]

Solutions

2.1 a value inside the interval is plausible; a value outside is not supported by the data.

2.2 yes, plausible (0.5 is inside the interval).

2.3 the data do not support that claim.

3.1 B.

3.2 B.

3.3 (a) yes (the whole interval is above 0.5). (b) no. (c) 95%.

Name:

AP Statistics: 2022

A survey conducted by a national research center asked a random sample of 920 teenagers in the United States how often they use a video streaming service. From the sample, 59% answered that they use a video streaming service every day.

(a) Construct and interpret a 95% confidence interval for the proportion of all teenagers in the United States who would respond that they use a video streaming service every day.

GO ON TO THE NEXT PAGE.

Continue your response to **QUESTION 4** on this page.

- (b) Based on the confidence interval in part (a), do the sample data provide convincing statistical evidence that the proportion of all teenagers in the United States who would respond that they use a video streaming service every day is not 0.5? Justify your answer.

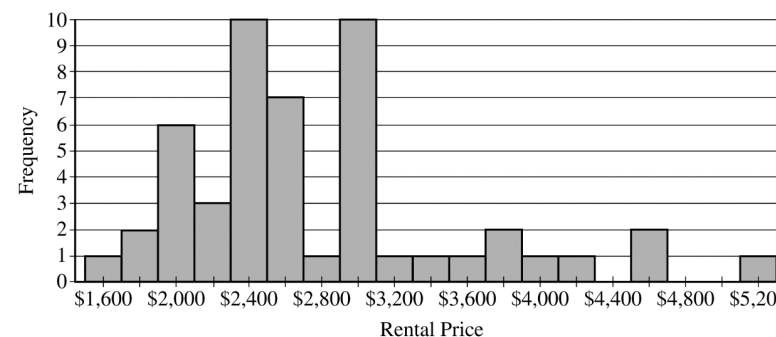
GO ON TO THE NEXT PAGE.**STATISTICS****SECTION II****Part B****Question 6****Spend about 25 minutes on this part of the exam.****Percent of Section II score—25**

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Emma is moving to a large city and is investigating typical monthly rental prices of available one-bedroom apartments. She obtained a random sample of rental prices for 50 one-bedroom apartments taken from a Web site where people voluntarily list available apartments.

- (a) Describe the population for which it is appropriate for Emma to generalize the results from her sample.

The distribution of the 50 rental prices of the available apartments is shown in the following histogram.



- (b) Emma wants to estimate the typical rental price of a one-bedroom apartment in the city. Based on the distribution shown, what is a disadvantage of using the mean rather than the median as an estimate of the typical rental price?
- (c) Instead of using the sample median as the point estimate for the population median, Emma wants to use an interval estimate. However, computing an interval estimate requires knowing the sampling distribution of the sample median for samples of size 50. Emma has one point, her sample median, in that sampling distribution. Using information about rental prices that are available on the Web site, describe how someone could develop a theoretical sampling distribution of the sample median for samples of size 50.

Because Emma does not have the resources to develop the theoretical sampling distribution, she estimates the sampling distribution of the sample median using a process called bootstrapping. In the bootstrapping process, a computer program performs the following steps.

- Take a random sample, with replacement, of size 50 from the original sample.
- Calculate and record the median of the sample.
- Repeat the process to obtain a total of 15,000 medians.

Emma ran the bootstrap process, and the following frequency table is the bootstrap distribution showing her results of generating 15,000 medians.

Bootstrap Distribution of Medians					
Median	Frequency	Median	Frequency	Median	Frequency
2,345	1	2,585	1	2,825	247
2,390	13	2,587.5	171	2,837.5	7
2,395	18	2,600	22	2,847.5	1
2,400	56	2,612.5	1,190	2,872.5	317
2,445	4	2,625	174	2,885	10
2,447.5	56	2,672.5	5	2,950	700
2,450	55	2,675	1,924	2,962.5	93
2,475	3	2,687.5	1,341	2,972.5	6
2,495	66	2,700	2,825	2,975	65
2,497.5	136	2,735	35	2,985	12
2,500	1,899	2,747.5	619	2,987.5	1
2,522.5	2	2,750	2	2,995	6
2,525	945	2,795	278	3,000	2
2,550	1,673	2,812.5	16	3,062.5	3

The bootstrap distribution provides an approximation of the sampling distribution of the sample median. A confidence interval for the median can be constructed using a percentage of the values in the middle of the bootstrap distribution.

- (d) Use the frequency table to find the following.
- Value of the 5th percentile:
 - Value of the 95th percentile:
- (e) Find the percentage of bootstrap medians in the table that are equal to or between the values found in part (d).
- (f) Use your values from parts (d) and (e) to construct and interpret a confidence interval for the median rental price.

STOP

END OF EXAM

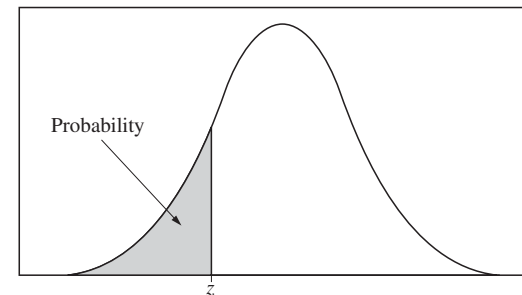


Table entry for z is the probability lying below z .

Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

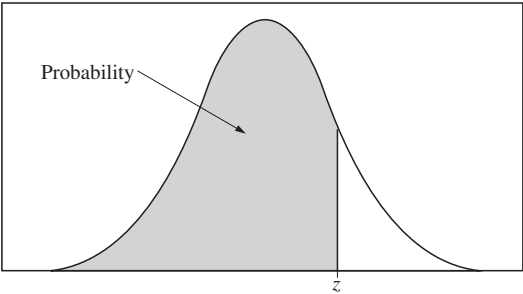


Table entry for z is the probability lying below z .

Table A (Continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

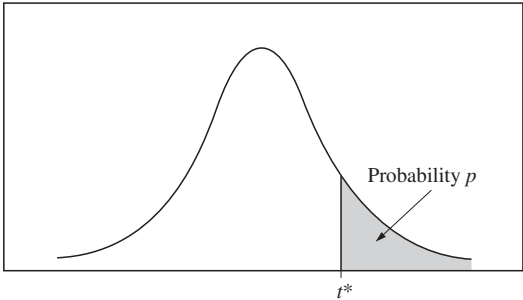
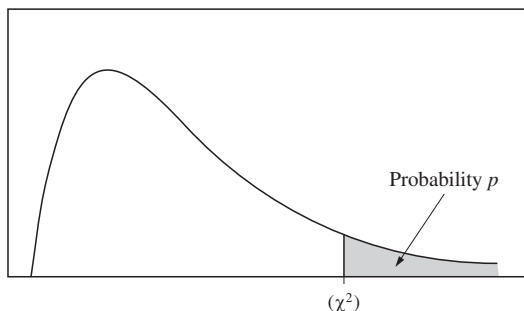


Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^* .

Table B t distribution critical values

df	Tail probability p									
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807
Confidence level C										
50% 60% 70% 80% 90% 95% 96% 98% 99% 99.5% 99.8% 99.9%										

Table entry for p is the point (χ^2) with probability p lying above it.

Table C χ^2 critical values

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

2. The manager of a local fast-food restaurant is concerned about customers who ask for a water cup when placing an order but fill the cup with a soft drink from the beverage fountain instead of filling the cup with water. The manager selected a random sample of 80 customers who asked for a water cup when placing an order and found that 23 of those customers filled the cup with a soft drink from the beverage fountain.
 - (a) Construct and interpret a 95 percent confidence interval for the proportion of all customers who, having asked for a water cup when placing an order, will fill the cup with a soft drink from the beverage fountain.
 - (b) The manager estimates that each customer who asks for a water cup but fills it with a soft drink costs the restaurant \$0.25. Suppose that in the month of June 3,000 customers ask for a water cup when placing an order. Use the confidence interval constructed in part (a) to give an interval estimate for the cost to the restaurant for the month of June from the customers who ask for a water cup but fill the cup with a soft drink.

2. To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

The owner selected a random sample of bills and found that only 15 percent of them received discounts. A confidence interval for p , the proportion of bills that will receive a discount in the long run, is 0.15 ± 0.06 . All conditions for inference were met.

- (a) Consider the confidence interval 0.15 ± 0.06 .
 - (i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.
 - (ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2? Justify your answer.

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

- (b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for p with the same confidence level as that of the original interval.
- (c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

6.4 Setting Up a Test for a Population Proportion

Name: _____ Class: _____ Date: _____ Total: 9 marks

KEY WORDS null 零假设 • test statistic 检验统计量 • null hypothesis 原假设
• alternative hypotheses 备择假设

Objective

Build the skills to answer exam questions on **setting up a test for a population proportion**.

You must be able to:

- state the **null** 零假设 and **alternative hypotheses** 备择假设
- write the test statistic $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
- state the conditions for the test

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ Hypotheses

- Null** $H_0 : p = p_0$ (a specific value — “no effect”).
- Alternative** $H_a : p \neq p_0$ (or $< p_0$, or $> p_0$).

■ Test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

The standard error uses the **null value** p_0 . Conditions: **random, independent** (10% rule), and **large counts** ($np_0 \geq 10$, $n(1-p_0) \geq 10$).

2 Practice

2.1 State the null hypothesis for a test about a proportion. [1]

2.2 Write the test-statistic formula for a proportion. [1]

2.3 State one condition needed to run the test. [1]

3 Exam-style questions

3.1 The null hypothesis H_0 usually states that [1]

- A** there is an effect
- B** there is no effect (a specific value)
- C** the sample is biased
- D** nothing at all

3.2 The test statistic for a proportion computes the standard error using [1]

- A** $\hat{p}$
- B** p_0 (the null value)
- C** zero
- D** the sample size only

3.3 A test checks whether $p \neq 0.5$, using $\hat{p} = 0.6$ and $n = 100$.

(a) Write H_0 . [1]

(b) Write H_a . [1]

(c) Compute the test statistic z . [2]

Solutions

2.1 $H_0 : p = p_0$ (a specific value).

$$\mathbf{2.2} \ z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}.$$

2.3 any one of: random sample, independence (10% rule), large counts.

3.1 B.

3.2 B.

$$\mathbf{3.3} \text{ (a) } H_0 : p = 0.5. \text{ (b) } H_a : p \neq 0.5. \text{ (c) } z = \frac{0.6 - 0.5}{\sqrt{0.5(0.5)/100}} = \frac{0.1}{0.05} = 2.0.$$

4. A software application (app) lets users enter questions to receive answers in the form of images, texts, or videos. Research indicates that 22 percent of high school students in Country W use the app to help them with their homework at least once per week. Karen is an AP Statistics student in Country W at a high school that has more than 2,000 students. She believes the proportion of all students at her school who use the app to help them with their homework at least once per week is greater than the proportion for her country. To investigate her belief, she took a simple random sample of 130 students from her school and found that 38 of the sampled students use the app to help them with their homework at least once per week.

Is there convincing statistical evidence, at a 0.05 significance level, to support Karen's belief? Justify your answer with the appropriate inference procedure.

4. The manager of a large company that sells pet supplies online wants to increase sales by encouraging repeat purchases. The manager believes that if past customers are offered \$10 off their next purchase, more than 40 percent of them will place an order. To investigate the belief, 90 customers who placed an order in the past year are selected at random. Each of the selected customers is sent an e-mail with a coupon for \$10 off the next purchase if the order is placed within 30 days. Of those who receive the coupon, 38 place an order.
- (a) Is there convincing statistical evidence, at the significance level of $\alpha = 0.05$, that the manager's belief is correct? Complete the appropriate inference procedure to support your answer.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.

(b) Based on your conclusion from part (a), which of the two errors, Type I or Type II, could have been made? Interpret the consequence of the error in context.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.

6.5 Interpreting p-Values

Name: _____ Class: _____ Date: _____ Total: 8 marks

KEY WORDS p-value P 值 • null hypothesis 原假设

Objective

Build the skills to answer exam questions on **interpreting p-values**.

You must be able to:

- define a **p-value** as the probability of a result at least as extreme, **assuming H_0 is true**
- explain that a **small** p-value is evidence **against H_0**

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ What a p-value is

The **p-value** is the probability of getting a sample result **as extreme or more extreme** than the one observed, **assuming the null hypothesis is true**.

■ Interpreting it

- A **small** p-value means the observed data would be unlikely if H_0 were true → **evidence against H_0** .
- A **large** p-value means the data are consistent with H_0 .

2 Practice

2.1 Define a p-value. [1]

2.2 State what a small p-value suggests. [1]

2.3 State the assumption made when computing a p-value. [1]

3 Exam-style questions

3.1 A p-value is computed assuming [1]

- **A** H_a is true
 - **B** H_0 is true
 - **C** the sample is biased
 - **D** nothing
-

3.2 A small p-value gives evidence [1]

- **A** for H_0
 - **B** against H_0
 - **C** of bias
 - **D** of nothing
-

3.3 A test gives a p-value of 0.02.

(a) State what this value measures. [1]

(b) State whether it is strong or weak evidence against H_0 . [1]

(c) State the assumption behind it. [1]

Solutions

2.1 the probability of a result at least as extreme as the one observed, assuming H_0 is true.

2.2 evidence against H_0 .

2.3 that the null hypothesis is true.

3.1 B.

3.2 B.

3.3 (a) the probability of a result this extreme if H_0 is true. (b) strong. (c) that H_0 is true.

6.6 Concluding a Test for a Population Proportion

Name: _____ Class: _____ Date: _____ Total: 8 marks

KEY WORDS significance level 显著性水平 • p-value P 值

Objective

Build the skills to answer exam questions on **concluding a test for a proportion**.

You must be able to:

- compare the **p-value** to the **significance level** 显著性水平 α
- decide to **reject** or **fail to reject** H_0
- state a conclusion in context

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ The decision rule

Compare the p-value to α (often 0.05):

- if p-value $< \alpha \rightarrow$ **reject** H_0 (the result is statistically significant);
- if p-value $\geq \alpha \rightarrow$ **fail to reject** H_0 .

■ In context

"Reject H_0 " means there is convincing evidence for H_a ; "fail to reject" means there is **not** enough evidence — never "accept H_0 ".

2 Practice

2.1 State the decision rule using the p-value and α . [1]

2.2 A test has a p-value of 0.03 and $\alpha = 0.05$. State the decision. [1]

2.3 State what "fail to reject H_0 " means. [1]

3 Exam-style questions

3.1 If the p-value is less than α , you [1]

- **A** reject H_0
- **B** accept H_0
- **C** fail to reject H_0
- **D** do nothing

3.2 A test with p-value 0.10 and $\alpha = 0.05$ leads to [1]

- **A** rejecting H_0
- **B** failing to reject H_0
- **C** an error
- **D** bias

3.3 A test of $p \neq 0.5$ gives a p-value of 0.01 at $\alpha = 0.05$.

(a) State the decision. [1]

(b) State whether the result is statistically significant. [1]

(c) State the conclusion in context. [1]

Solutions

2.1 reject H_0 if the p-value $< \alpha$; otherwise fail to reject H_0 .

2.2 reject H_0 ($0.03 < 0.05$).

2.3 there is not enough evidence to reject H_0 .

3.1 A.

3.2 B.

3.3 (a) reject H_0 . (b) yes. (c) there is convincing evidence that $p \neq 0.5$.

Name:

AP Statistics: 2025

- 4.** A software application (app) lets users enter questions to receive answers in the form of images, texts, or videos. Research indicates that 22 percent of high school students in Country W use the app to help them with their homework at least once per week. Karen is an AP Statistics student in Country W at a high school that has more than 2,000 students. She believes the proportion of all students at her school who use the app to help them with their homework at least once per week is greater than the proportion for her country. To investigate her belief, she took a simple random sample of 130 students from her school and found that 38 of the sampled students use the app to help them with their homework at least once per week.

Is there convincing statistical evidence, at a 0.05 significance level, to support Karen's belief? Justify your answer with the appropriate inference procedure.

4. The manager of a large company that sells pet supplies online wants to increase sales by encouraging repeat purchases. The manager believes that if past customers are offered \$10 off their next purchase, more than 40 percent of them will place an order. To investigate the belief, 90 customers who placed an order in the past year are selected at random. Each of the selected customers is sent an e-mail with a coupon for \$10 off the next purchase if the order is placed within 30 days. Of those who receive the coupon, 38 place an order.

(a) Is there convincing statistical evidence, at the significance level of $\alpha = 0.05$, that the manager's belief is correct? Complete the appropriate inference procedure to support your answer.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.

- (b) Based on your conclusion from part (a), which of the two errors, Type I or Type II, could have been made? Interpret the consequence of the error in context.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.

6.7 Potential Errors When Performing Tests

Name: _____ Class: _____ Date: _____ **Total: 8 marks**

KEY WORDS type i error 第一类错误 • type ii error 第二类错误 • power 功效

Objective

Build the skills to answer exam questions on **potential errors in tests**.

You must be able to:

- describe a **Type I error** 第一类错误 (rejecting a true H_0)
- describe a **Type II error** 第二类错误 (failing to reject a false H_0)
- describe the **power** 功效 of a test

1 Worked examples

Study these first. Each one shows the method for a question type used later.

- **Two kinds of error**
- **Type I** — reject H_0 when it is actually **true** (a false alarm). Its probability is α .
 - **Type II** — fail to reject H_0 when it is actually **false** (a missed effect). Its probability is β .
- **Power**
- Power** = $1 - \beta$ — the probability of correctly detecting a real effect (rejecting a false H_0).

2 Practice

2.1 Define a Type I error. [1]

2.2 Define a Type II error. [1]

2.3 State what the power of a test measures. [1]

3 Exam-style questions

3.1 A Type I error is [1]

- **A** rejecting a true H_0
- **B** failing to reject a false H_0
- **C** a correct decision
- **D** a form of bias

3.2 The probability of a Type I error equals [1]

- **A** β
- **B** α
- **C** the power
- **D** 0

3.3 A drug that does **not** work is declared effective by a test.

- (a) Name this type of error. [1]
- (b) Name the error of missing a drug that **does** work. [1]
- (c) Name the probability of correctly detecting a real effect. [1]

Solutions

2.1 rejecting H_0 when it is actually true.

2.2 failing to reject H_0 when it is actually false.

2.3 the probability of correctly rejecting a false H_0 (detecting a real effect).

3.1 A.

3.2 B.

3.3 (a) a Type I error. (b) a Type II error. (c) power.

Name:

AP Statistics: 2025

5. According to a 2017 national survey in Country B, the mean number of bedrooms in newly built houses was 2.9. Rodney, a researcher, believes the mean number of bedrooms in newly built houses in the country was different in 2024 than it was in 2017. To investigate his belief, he took a large random sample of newly built houses in Country B in 2024 and recorded the number of bedrooms in each house. The distribution of the number of bedrooms for the sampled houses is summarized in the table.

Distribution of the Number of Bedrooms for the Houses Sampled in 2024

Number of Bedrooms	1	2	3	4	5	6
Proportion of Houses	0.12	0.22	0.28	0.22	0.14	0.02

A.

- A house from the sample will be selected at random. What is the probability that the house had fewer than 3 bedrooms? Show your work.
- What is the mean number of bedrooms for the sample of newly built houses in 2024? Show your work.

B. Rodney will use a one-sample t -test for a population mean to test his belief.

- In the context of Rodney's investigation, state the hypotheses for the test.
- Explain, in context, what a Type I error would be for Rodney's hypothesis test.

C. A different researcher, Keisha, suggests using a confidence interval to investigate whether the mean number of bedrooms in newly built houses in 2024 in Country B was different from 2.9.

Assume the conditions for inference have been met. Using Rodney's data, Keisha calculated a one-sample 97 percent confidence interval to estimate the population mean as $(3.01, 3.19)$. Based on the confidence interval, what conclusion can be made for Rodney's hypothesis test in part B at $\alpha = 0.03$? Justify your answer.

4. The manager of a large company that sells pet supplies online wants to increase sales by encouraging repeat purchases. The manager believes that if past customers are offered \$10 off their next purchase, more than 40 percent of them will place an order. To investigate the belief, 90 customers who placed an order in the past year are selected at random. Each of the selected customers is sent an e-mail with a coupon for \$10 off the next purchase if the order is placed within 30 days. Of those who receive the coupon, 38 place an order.

(a) Is there convincing statistical evidence, at the significance level of $\alpha = 0.05$, that the manager's belief is correct? Complete the appropriate inference procedure to support your answer.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.



- (b) Based on your conclusion from part (a), which of the two errors, Type I or Type II, could have been made? Interpret the consequence of the error in context.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.



STATISTICS
SECTION II

Part B

Question 6

Spend about 25 minutes on this part of the exam.

Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Systolic blood pressure is the amount of pressure that blood exerts on blood vessels while the heart is beating. The mean systolic blood pressure for people in the United States is reported to be 122 millimeters of mercury (mmHg) with a standard deviation of 15 mmHg.

The wellness department of a large corporation is investigating whether the mean systolic blood pressure of its employees is greater than the reported national mean. A random sample of 100 employees will be selected, the systolic blood pressure of each employee in the sample will be measured, and the sample mean will be calculated.

Let μ represent the mean systolic blood pressure of all employees at the corporation. Consider the following hypotheses.

$$H_0 : \mu = 122$$

$$H_a : \mu > 122$$

- (a) Describe a Type II error in the context of the hypothesis test.
- (b) Assume that σ , the standard deviation of the systolic blood pressure of all employees at the corporation, is 15 mmHg. If $\mu = 122$, the sampling distribution of $\bar{x}$ for samples of size 100 is approximately normal with a mean of 122 mmHg and a standard deviation of 1.5 mmHg. What values of the sample mean $\bar{x}$ would represent sufficient evidence to reject the null hypothesis at the significance level of $\alpha = 0.05$?

The actual mean systolic blood pressure of all employees at the corporation is 125 mmHg, not the hypothesized value of 122 mmHg, and the standard deviation is 15 mmHg.

- (c) Using the actual mean of 125 mmHg and the results from part (b), determine the probability that the null hypothesis will be rejected.
- (d) What statistical term is used for the probability found in part (c) ?
- (e) Suppose the size of the sample of employees to be selected is greater than 100. Would the probability of rejecting the null hypothesis be greater than, less than, or equal to the probability calculated in part (c) ? Explain your reasoning.

STOP

END OF EXAM

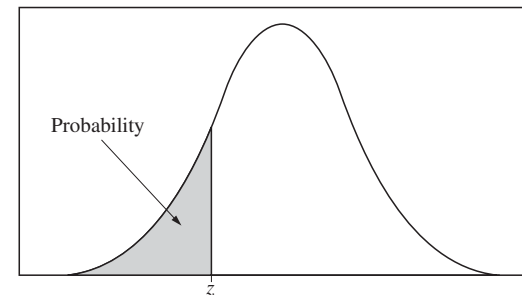


Table entry for z is the probability lying below z .

Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



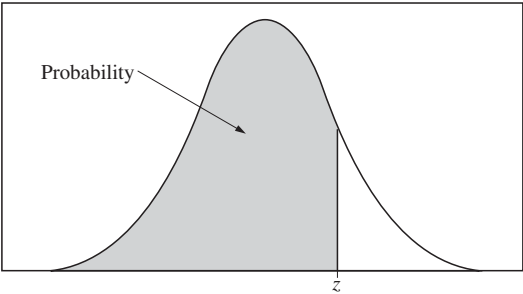


Table entry for z is the probability lying below z .

Table A (Continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

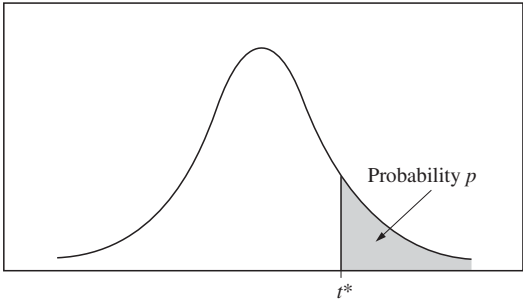


Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^* .

Table B t distribution critical values

df	Tail probability p									
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807

50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------

Confidence level C

Table entry for p is the point (χ^2) with probability p lying above it.

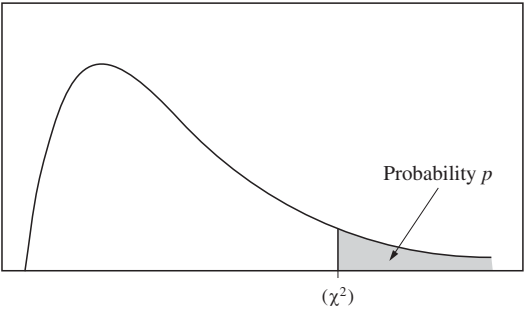


Table C χ^2 critical values

df	Tail probability p										
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4

AP STATISTICS • EXERCISE SHEET

6.8 Confidence Intervals for the Difference of Two Proportions

Name: _____ Class: _____ Date: _____ **Total: 8 marks**

KEY WORDS confidence interval 置信区间

Objective

Build the skills to answer exam questions on **a confidence interval for the difference of two proportions**.

You must be able to:

- use the point estimate $\hat{p}_1 - \hat{p}_2$
- build the interval $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$
- use the interval to judge whether the two proportions differ

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ The two-proportion interval

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}.$$

■ Judging a difference

If the interval **contains** 0, the two proportions may not differ (no significant difference). If it is **entirely above (or below)** 0, there is a significant difference.

2 Practice

2.1 Write the point estimate for the difference of two proportions. [1]

2.2 State how to judge whether the two proportions differ using the interval. [1]

2.3 For $\hat{p}_1 = 0.6$ and $\hat{p}_2 = 0.5$, find the point estimate of the difference. [1]

3 Exam-style questions

3.1 A confidence interval for $p_1 - p_2$ is centred on [1]

- **A** $\hat{p}_1 \cdot \hat{p}_2$
 - **B** $\hat{p}_1 - \hat{p}_2$
 - **C** 0
 - **D** $\hat{p}_1 + \hat{p}_2$
-

3.2 If a 95% CI for $p_1 - p_2$ contains 0, the two proportions [1]

- **A** definitely differ
 - **B** may not differ (no significant difference)
 - **C** are exactly equal
 - **D** are both zero
-

3.3 A 95% CI for $p_1 - p_2$ is (0.02, 0.10).

(a) State which proportion is larger. [1]

(b) State whether 0 is in the interval. [1]

(c) State whether there is a significant difference. [1]

Solutions

2.1 $\hat{p}_1 - \hat{p}_2$.

2.2 if the interval contains 0, there is no significant difference; if not, there is.

2.3 $0.6 - 0.5 = 0.1$.

3.1 **B**.

3.2 **B**.

3.3 (a) p_1 (the interval is positive). (b) no. (c) yes.

5. A polling agency showed the following two statements to a random sample of 1,048 adults in the United States.

Environment statement: Protection of the environment should be given priority over economic growth.

Economy statement: Economic growth should be given priority over protection of the environment.

The order in which the statements were shown was randomly selected for each person in the sample. After reading the statements, each person was asked to choose the statement that was most consistent with his or her opinion. The results are shown in the table.

	Environment Statement	Economy Statement	No Preference
Percent of sample	58%	37%	5%

- (a) Assume the conditions for inference have been met. Construct and interpret a 95 percent confidence interval for the proportion of all adults in the United States who would have chosen the economy statement.
- (b) One of the conditions for inference that was met is that the number who chose the economy statement and the number who did not choose the economy statement are both greater than 10. Explain why it is necessary to satisfy that condition.
- (c) A suggestion was made to use a two-sample z -interval for a difference between proportions to investigate whether the difference in proportions between adults in the United States who would have chosen the environment statement and adults in the United States who would have chosen the economy statement is statistically significant. Is the two-sample z -interval for a difference between proportions an appropriate procedure to investigate the difference? Justify your answer.

AP STATISTICS • EXERCISE SHEET

6.9 Justifying a Claim Based on a Confidence Interval for a Difference of Population Proportions

Name: _____ Class: _____ Date: _____ Total: 8 marks

KEY WORDS confidence interval 置信区间

Objective

Build the skills to answer exam questions on **justifying a claim from a confidence interval for a difference of proportions**.

You must be able to:

- use a confidence interval for $p_1 - p_2$ to judge a claim
- recognise that an interval containing 0 means **no significant difference**
- interpret an interval entirely above or below 0

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ Judging a difference from the interval

A confidence interval for $p_1 - p_2$ gives the plausible values of the difference:

- if the interval **contains** 0, "no difference" ($p_1 = p_2$) is plausible;
- if the interval is entirely **positive**, $p_1 > p_2$; entirely **negative**, $p_1 < p_2$.

■ Reading a worked interval

A study reports the difference $p_1 - p_2$ as the 95% interval $(-0.02, 0.11)$. Because it **contains** 0, a zero difference is plausible – the data do **not** show a significant difference. Had it been $(0.03, 0.11)$, entirely positive, we would conclude $p_1 > p_2$.

2 Practice

2.1 State how a CI for $p_1 - p_2$ can justify a claim of no difference. [1]

2.2 A 95% CI for $p_1 - p_2$ is $(-0.03, 0.05)$. State whether "no difference" is plausible. [1]

2.3 State what a fully positive interval implies. [1]

3 Exam-style questions

3.1 A claim of "no difference" ($p_1 = p_2$) is plausible if the CI for $p_1 - p_2$ contains [1]

- **A** 1
- **B** 0
- **C** $\hat{p}_1$
- **D** the mean

3.2 A 95% CI for $p_1 - p_2$ is $(0.04, 0.12)$. The claim $p_1 = p_2$ is [1]

- **A** supported
- **B** not supported
- **C** certainly true
- **D** unknown

3.3 A 95% CI for the difference in approval between two groups is $(-0.02, 0.08)$.

(a) State whether 0 is in the interval. [1]

(b) State whether the groups differ significantly. [1]

(c) State the plausible conclusion. [1]

Solutions

2.1 if the interval contains 0, "no difference" is plausible.

2.2 yes, plausible (0 is inside the interval).

2.3 that $p_1 > p_2$ (a real difference).

3.1 B.

3.2 B.

3.3 (a) yes. (b) no. (c) there is no significant difference between the groups.

6.10 Setting Up a Test for the Difference of Two Population Proportions

Name: _____ Class: _____ Date: _____ Total: 8 marks

KEY WORDS pooled (combined) proportion 合并比例 • null hypothesis 原假设
• test statistic 检验统计量

Objective

Build the skills to answer exam questions on **setting up a test for a difference of two proportions**.

You must be able to:

- state the hypotheses $H_0 : p_1 = p_2$ and H_a
- explain why a **pooled (combined) proportion** 合并比例 is used
- write the pooled proportion $\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ Hypotheses

- Null** $H_0 : p_1 = p_2$ (no difference).
- Alternative** $H_a : p_1 \neq p_2$ (or one-sided).

■ The pooled proportion

Because H_0 assumes the proportions are **equal**, we combine both samples to estimate that common value:

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2},$$

and use $\hat{p}_c$ in the standard error of the test statistic.

2 Practice

2.1 State the null hypothesis for a two-proportion test. [1]

2.2 State why a pooled proportion is used in the test. [1]

2.3 Write the formula for the pooled proportion. [1]

3 Exam-style questions

3.1 The null hypothesis in a two-proportion test is [1]

- A $p_1 \neq p_2$
- B $p_1 = p_2$
- C $p_1 = 0$
- D $\hat{p}_1 = \hat{p}_2$

3.2 The pooled proportion combines [1]

- A only group 1
- B both groups' successes and sample sizes
- C the two means
- D nothing

3.3 Group 1 has 30 successes in 100; group 2 has 20 in 100.

(a) Find $\hat{p}_1$. [1]

(b) Find $\hat{p}_2$. [1]

(c) Find the pooled proportion $\hat{p}_c$. [1]

Solutions

2.1 $H_0 : p_1 = p_2$.

2.2 H_0 assumes the proportions are equal, so both samples estimate the same common value.

2.3 $\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$.

3.1 B.

3.2 B.

3.3 (a) 0.30. (b) 0.20. (c) $\hat{p}_c = \frac{30 + 20}{100 + 100} = 0.25$.

Name:

AP Statistics: 2024

1. Descriptive Statistics

$$\bar{x} = \frac{1}{n} \sum x_i = \frac{\sum x_i}{n}$$

$$\hat{y} = a + bx$$

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

$$s_x = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$\bar{y} = a + b\bar{x}$$

$$b = r \frac{s_y}{s_x}$$

4. Tumbleweed, commonly found in the western United States, is the dried structure of certain plants that are blown by the wind. Kochia, a type of plant that turns into tumbleweed at the end of the summer, is a problem for farmers because it takes nutrients away from soil that would otherwise go to more beneficial plants. Scientists are concerned that kochia plants are becoming resistant to the most commonly used herbicide, glyphosate. In 2014, 19.7 percent of 61 randomly selected kochia plants were resistant to glyphosate. In 2017, 38.5 percent of 52 randomly selected kochia plants were resistant to glyphosate. Do the data provide convincing statistical evidence, at the level of $\alpha = 0.05$, that there has been an increase in the proportion of all kochia plants that are resistant to glyphosate?

4. A researcher conducted a medical study to investigate whether taking a low-dose aspirin reduces the chance of developing colon cancer. As part of the study, 1,000 adult volunteers were randomly assigned to one of two groups. Half of the volunteers were assigned to the experimental group that took a low-dose aspirin each day, and the other half were assigned to the control group that took a placebo each day. At the end of six years, 15 of the people who took the low-dose aspirin had developed colon cancer and 26 of the people who took the placebo had developed colon cancer. At the significance level $\alpha = 0.05$, do the data provide convincing statistical evidence that taking a low-dose aspirin each day would reduce the chance of developing colon cancer among all people similar to the volunteers?

6.11 Carrying Out a Test for the Difference of Two Population Proportions

Name: _____ Class: _____ Date: _____ Total: 9 marks

KEY WORDS significance level 显著性水平 • p-value P 值 • test statistic 检验统计量

Objective

Build the skills to answer exam questions on **carrying out a test for a difference of two proportions**.

You must be able to:

- compute the test statistic z using the pooled standard error
- find the **p-value** and compare it to α
- state a conclusion in context

1 Worked examples

Study these first. Each one shows the method for a question type used later.

■ The full procedure

1. State $H_0 : p_1 = p_2$ and H_a ; check conditions.
2. Compute $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_c(1 - \hat{p}_c)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$.
3. Find the **p-value** from z .
4. Compare to α : reject H_0 if p-value $< \alpha$, else fail to reject.
5. State the conclusion **in context**.

■ A worked test

$\hat{p}_1 = 0.60$ ($n_1 = 50$), $\hat{p}_2 = 0.45$ ($n_2 = 60$). Pooled $\hat{p}_c = \frac{30+27}{110} = 0.518$. Then

$$z = \frac{0.60 - 0.45}{\sqrt{0.518(0.482)\left(\frac{1}{50} + \frac{1}{60}\right)}} \approx \frac{0.15}{0.096} \approx 1.57,$$

a two-sided p-value $\approx 0.12 > 0.05$ – fail to reject H_0 .

2 Practice

2.1 State the steps to carry out a two-proportion test. [2]

2.2 A two-proportion test gives $z = 2.5$. State whether this is likely significant at $\alpha = 0.05$. [1]

2.3 State what to compare the p-value to. [1]

3 Exam-style questions

3.1 After computing z , you find the [1]

- A slope
- B p-value
- C mean
- D residual

3.2 A two-proportion test rejects H_0 when the p-value is [1]

- A above α
- B below α
- C exactly zero
- D exactly one

3.3 A two-proportion test gives a p-value of 0.03 at $\alpha = 0.05$.

- (a) State the decision. [1]
(b) State whether the two proportions differ significantly. [1]
(c) State the conclusion in context. [1]

Solutions

2.1 state hypotheses and check conditions; compute z ; find the p-value; compare to α and conclude in context.

2.2 yes — $z = 2.5$ gives a small p-value, likely below 0.05.

2.3 the significance level α .

3.1 B.

3.2 B.

3.3 (a) reject H_0 . (b) yes. (c) there is convincing evidence that the two proportions differ.

I. Descriptive Statistics

$$\bar{x} = \frac{1}{n} \sum x_i = \frac{\sum x_i}{n}$$

$$\hat{y} = a + bx$$

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

$$s_x = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$\bar{y} = a + b\bar{x}$$

$$b = r \frac{s_y}{s_x}$$

4. Tumbleweed, commonly found in the western United States, is the dried structure of certain plants that are blown by the wind. Kochia, a type of plant that turns into tumbleweed at the end of the summer, is a problem for farmers because it takes nutrients away from soil that would otherwise go to more beneficial plants. Scientists are concerned that kochia plants are becoming resistant to the most commonly used herbicide, glyphosate. In 2014, 19.7 percent of 61 randomly selected kochia plants were resistant to glyphosate. In 2017, 38.5 percent of 52 randomly selected kochia plants were resistant to glyphosate. Do the data provide convincing statistical evidence, at the level of $\alpha = 0.05$, that there has been an increase in the proportion of all kochia plants that are resistant to glyphosate?

4. A researcher conducted a medical study to investigate whether taking a low-dose aspirin reduces the chance of developing colon cancer. As part of the study, 1,000 adult volunteers were randomly assigned to one of two groups. Half of the volunteers were assigned to the experimental group that took a low-dose aspirin each day, and the other half were assigned to the control group that took a placebo each day. At the end of six years, 15 of the people who took the low-dose aspirin had developed colon cancer and 26 of the people who took the placebo had developed colon cancer. At the significance level $\alpha = 0.05$, do the data provide convincing statistical evidence that taking a low-dose aspirin each day would reduce the chance of developing colon cancer among all people similar to the volunteers?