2018

AP® CollegeBoard

AP Statistics

Free-Response Questions

Formulas begin on page 3.
Questions begin on page 6.
Tables begin on page 12.

Formulas

(I) Descriptive Statistics

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

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

$$s_p = \sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum\left(x_i - \bar{x}\right)^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

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

$$b_1 = r\frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\dfrac{\sum\left(y_i - \hat{y}_i\right)^2}{n-2}}}{\sqrt{\sum\left(x_i - \bar{x}\right)^2}}$$

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

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

$$E(X) = \mu_x = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum \left(x_i - \mu_x\right)^2 p_i$$

If X has a binomial distribution with parameters n and p, then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_x = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

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

If $\bar{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then:

$$\mu_{\bar{x}} = \mu$$

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

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) • (standard deviation of statistic)

Single-Sample

Statistic	**Standard Deviation of Statistic**
Sample Mean	$\dfrac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\dfrac{p(1-p)}{n}}$

Two-Sample

Statistic	**Standard Deviation**
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Chi-square test statistic $= \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

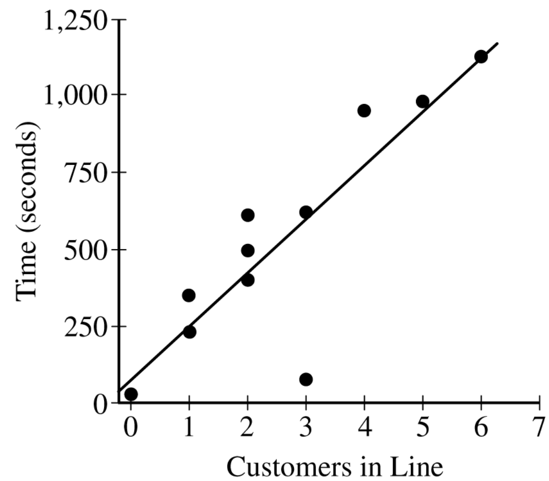
STATISTICS
SECTION II
Part A
Questions 1-5

Spend about 1 hour and 5 minutes on this part of the exam.

Percent of Section II score—75

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.

1. The manager of a grocery store selected a random sample of 11 customers to investigate the relationship between the number of customers in a checkout line and the time to finish checkout. As soon as the selected customer entered the end of a checkout line, data were collected on the number of customers in line who were in front of the selected customer and the time, in seconds, until the selected customer was finished with the checkout. The data are shown in the following scatterplot along with the corresponding least-squares regression line and computer output.



Predictor	Coef	SE Coef	T	P
Constant	72.95	110.36	0.66	0.525
Customers in line	174.40	35.06	4.97	0.001
S = 200.01		R-Sq = 73.33%	R-Sq (adj) = 70.37%	

(a) Identify and interpret in context the estimate of the intercept for the least-squares regression line.

(b) Identify and interpret in context the coefficient of determination, r^2.

(c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

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.

GO ON TO THE NEXT PAGE.

3. Approximately 3.5 percent of all children born in a certain region are from multiple births (that is, twins, triplets, etc.). Of the children born in the region who are from multiple births, 22 percent are left-handed. Of the children born in the region who are from single births, 11 percent are left-handed.

   (a) What is the probability that a randomly selected child born in the region is left-handed?

   (b) What is the probability that a randomly selected child born in the region is a child from a multiple birth, given that the child selected is left-handed?

   (c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

4. The anterior cruciate ligament (ACL) is one of the ligaments that help stabilize the knee. Surgery is often recommended if the ACL is completely torn, and recovery time from the surgery can be lengthy. A medical center developed a new surgical procedure designed to reduce the average recovery time from the surgery. To test the effectiveness of the new procedure, a study was conducted in which 210 patients needing surgery to repair a torn ACL were randomly assigned to receive either the standard procedure or the new procedure.

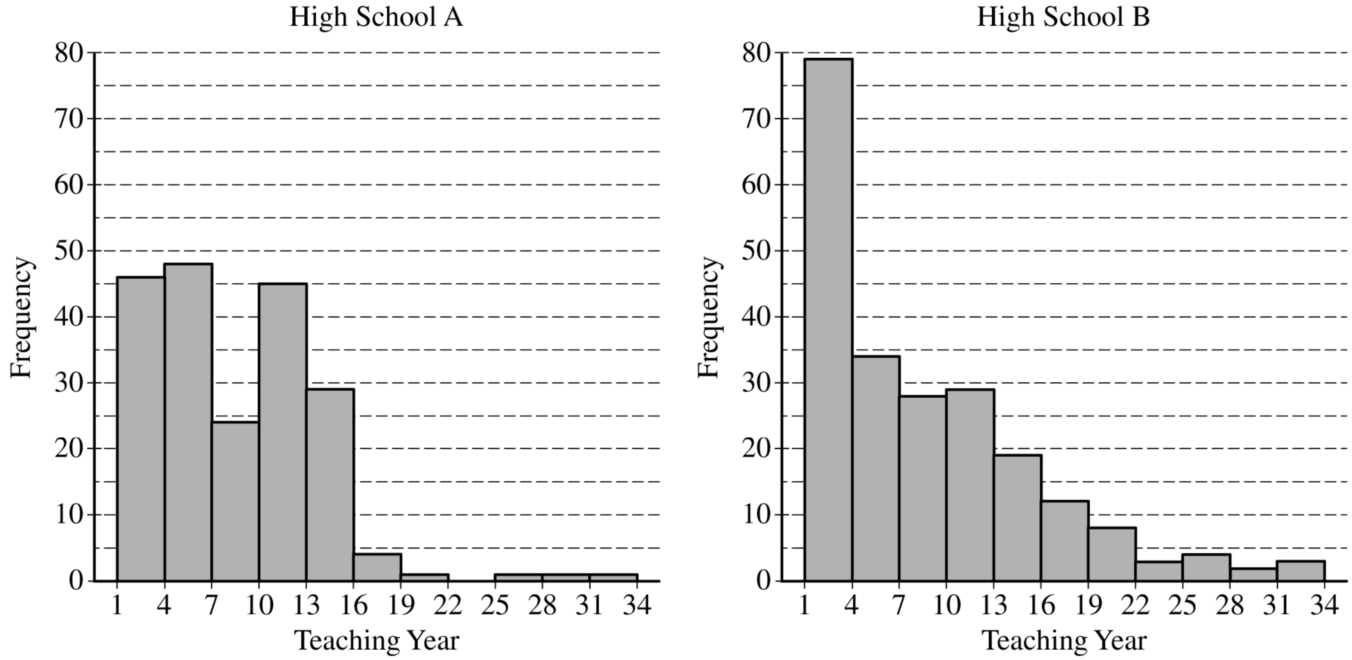
(a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

(b) Summary statistics on the recovery times from the surgery are shown in the table.

Type of Procedure	Sample Size	Mean Recovery Time (days)	Standard Deviation Recovery Time (days)
Standard	110	217	34
New	100	186	29

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

5. The following histograms summarize the teaching year for the teachers at two high schools, A and B.



Teaching year is recorded as an integer, with first-year teachers recorded as 1, second-year teachers recorded as 2, and so on. Both sets of data have a mean teaching year of 8.2, with data recorded from 200 teachers at High School A and 221 teachers at High School B. On the histograms, each interval represents possible integer values from the left endpoint up to but not including the right endpoint.

(a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

(b) An additional 18 teachers were not included with the data recorded from the 200 teachers at High School A. The mean teaching year of the 18 teachers is 2.5. What is the mean teaching year for all 218 teachers at High School A?

(c) The standard deviation of the teaching year for the 221 teachers at High School B is 7.2. If one teacher is selected at random from High School B, what is the probability that the teaching year for the selected teacher will be within 1 standard deviation of the mean of 8.2 ? Justify your answer.

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.

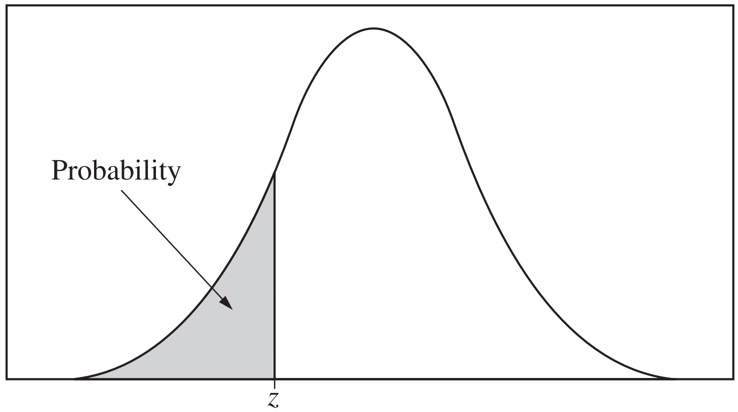


Table A Standard normal probabilities

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.

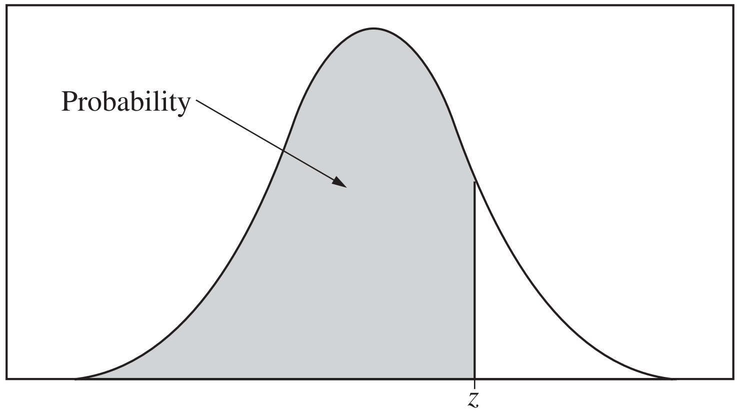


Table A (*Continued*)

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^*.

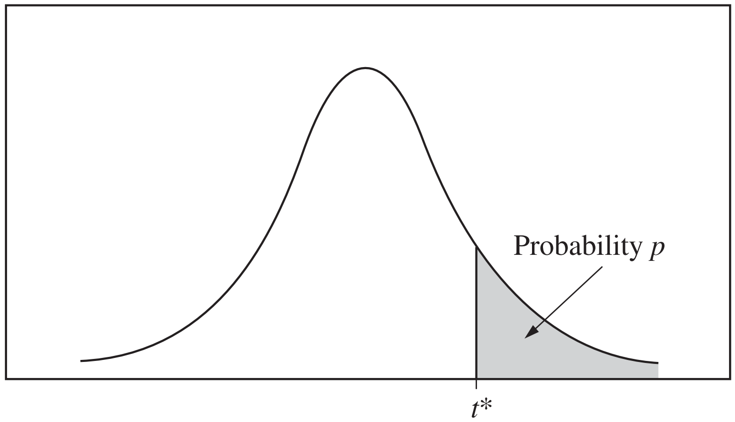


Table B ***t*** **distribution critical values**

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.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.

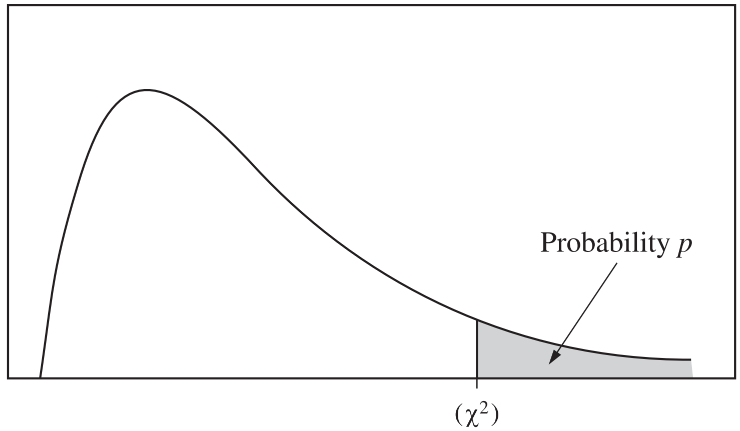


Table C χ^2 **critical values**

	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

2018

AP CollegeBoard

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free Response Question 1

- ☑ Scoring Guideline
- ☑ Student Samples
- ☑ Scoring Commentary

Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) identify various values in regression computer output; (2) interpret the intercept of a regression line in context; (3) interpret the coefficient of determination (r^2) in context; and (4) identify an outlier from a scatterplot.

Solution

Part (a):

The estimate of the intercept is 72.95. It is estimated that the average time to finish checkout if there are no other customers in line is 72.95 seconds.

Part (b):

The coefficient of determination is $r^2 = 73.33\%$. This value indicates that 73.33% of the variability in the times it takes customers to finish checkout, including time waiting in line, can be explained by knowing how many customers are in line in front of the selected customer.

Part (c):

The outlier is the point with $x = 3$ and y close to 0. This point is considered an outlier because the combination of x and y values differs from the pattern of the rest of the data. Specifically, the value of y (time to finish checkout) is much lower than would be expected when there are $x = 3$ customers in line in front of the selected customer, given the remaining data.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Correctly identifies 72.95 as the intercept.
2. Communicates the concept of a y-intercept in a context that includes both time and zero customers.
3. Indicates that the value of the intercept is a prediction by using language such as "predicted," "estimated," or "average" value of y.

Partially correct (P) if the response includes only two of the three components.

Incorrect (I) if the response includes at most one of the three components.

Question 1 (continued)

Notes:
- Regression equations (such as $\hat{y} = 72.95 + 174.40x$) cannot be used to satisfy identification of the intercept in component 1, unless the intercept is explicitly labeled as such.
- A regression equation cannot be used to satisfy component 3.
- Incorrect regression equations are treated as extraneous and do not affect the scoring of any component.
- A response that interprets 72.95 as a slope does not satisfy components 1 or 2.

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:
1. Correctly identifies 73.33% as the coefficient of determination.
2. Provides a correct (possibly generic) interpretation of r^2.
3. Interpretation includes context.

Partially correct (P) if the response satisfies only two of the three components;
OR
if the response satisfies the three components, but reverses the roles of number of customers in line and time to finish checkout in the interpretation.

Incorrect (I) if the response satisfies at most one of the three components.

Notes:
- In component 2 the correct interpretation of the coefficient of determination can take any of several equivalent forms, such as:
  - The percent variability in y that is attributed to the linear relationship between y and x or between x and y.
  - The proportion of the total variability in the dependent variable y that is explained by the independent variable x.
  - The proportion of variation in y that is accounted for by the linear model.
  - The proportionate reduction of total variation of the y values that is associated with the use of the independent variable x.
  - The proportionate reduction in the sum of the squares of vertical deviations obtained by using the least-squares line instead of the naïve prediction of $\overline{y}$.
- In component 2 common *incorrect* interpretations of the coefficient of determination include:
  - The percent variability in the *predicted* y values that is explained by the linear relationship between y and x.
  - The percent variability in the *data* that is explained by the linear relationship between y and x.
  - The percent variability that is explained by the linear relationship between y and x.
  - The percent variability in y that is *on average* explained by the linear relationship between y and x.
- For component 3 context must include mention of time or customers.

Question 1 (continued)

Part (c) is scored as follows:

Essentially correct (E) if the response satisfies the following two components:

1. Correctly identifies the outlier.
2. Describes an unusual feature of the identified scatter plot point, relative to the remaining data points, that is sufficient to identify it as the outlier. Examples include:
   - The combination of x and y values is unusual compared to the other points.
   - The value of y is much lower than would be expected (or predicted), given the remaining data.
   - The residual for the point is unusually large relative to the other residuals.

Partially correct (P) if the response satisfies component 1 but does not satisfy component 2.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- In the absence of any point being circled on the graph, component 1 can still be satisfied by explicitly referring to the coordinates of the outlier. Valid coordinates for outlier identification must specify an x value of 3 and a y value that is strictly between 0 and 250.
- A response that does not make a comparison to the remaining data points, such as stating the outlier has a large residual or is nowhere near the regression line, does not satisfy component 2.
- A response that makes a comparison to the remaining data points based upon an unusual feature that is *insufficient* for outlier identification, such as stating the point is the only point with that particular y value, does not satisfy component 2.
- In the absence of explicit numerical calculation, a response that appeals to the influence that the outlier has on the regression coefficient estimates or on the sample correlation coefficient does not satisfy component 2.

Question 1 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

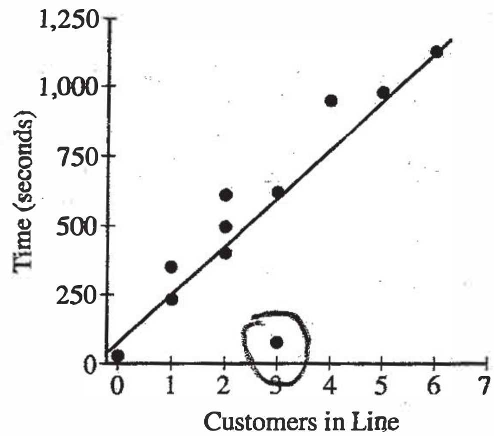
No parts essentially correct and two parts partially correct

1A-1

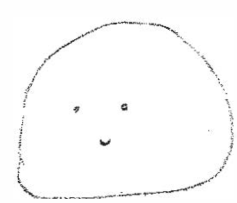
STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 1 hour and 5 minutes on this part of the exam.
Percent of Section II score—75

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.

1. The manager of a grocery store selected a random sample of 11 customers to investigate the relationship between the number of customers in a checkout line and the time to finish checkout. As soon as the selected customer entered ~~the end of a checkout line, data were collected on the number of customers~~ in line who were in front of the selected customer and the time, in seconds, until the selected customer was finished with the checkout. The data are shown in the following scatterplot along with the corresponding least-squares regression ~~line and computer~~ output.



Predictor	Coef	SE Coef	T	P
Constant	72.95	110.36	0.66	0.525
Customers in line	174.40	35.06	4.97	0.001
S = 200.01		R-Sq = 73.33%	R-Sq (adj) = 70.37%	



Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

1A2

(a) Identify and interpret in context the estimate of the intercept for the least-squares regression line.

The y intercept is 72.95 seconds. If there are no people in front of the customer, we would expect them to be finished with checkout in 72.95 seconds.

(b) Identify and interpret in context the coefficient of determination, r^2.

r^2 is 73.33%. 73.33% of the variation in checkout time is accounted for by the linear relationship between customers in line and checkout time.

(c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

That point is an outlier because it does not follow the pattern the rest of the data points follow and is very far from the rest of the data.

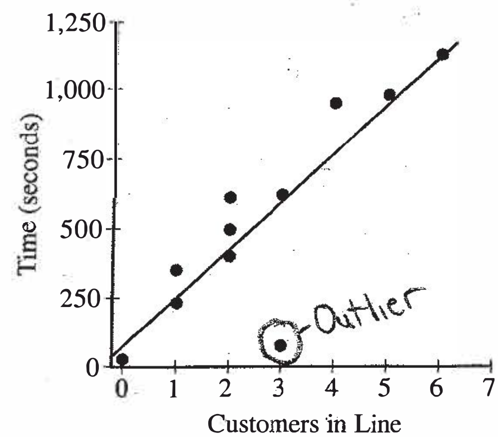
GO ON TO THE NEXT PAGE.

1B 1

STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 1 hour and 5 minutes on this part of the exam.
Percent of Section II score—75

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.

1. The manager of a grocery store selected a random sample of 11 customers to investigate the relationship between the number of customers in a checkout line and the time to finish checkout. As soon as the selected customer entered the end of a checkout line, data were collected on the number of customers in line who were in front of the selected customer and the time, in seconds, until the selected customer was finished with the checkout. The data are shown in the following scatterplot along with the corresponding least-squares regression line and computer output.



Predictor	Coef	SE Coef	T	P
Constant	72.95	110.36	0.66	0.525
Customers in line	174.40	35.06	4.97	0.001
S = 200.01		R-Sq = 73.33%	R-Sq (adj) = 70.37%	

GO ON TO THE NEXT PAGE.

1B 2

(a) Identify and interpret in context the estimate of the intercept for the least-squares regression line.

The estimate of the intercept, which is 72.95 seconds, means that if there are no customers in the line, the predicted time to finish checkout is 72.95 seconds.

(b) Identify and interpret in context the coefficient of determination, r^2.

The r^2 value of 73.33% means that about 73.33% of the variation of time to finish checkout, y, can be explained by the least-squares regression line of customers in line, x, and time to finish checkout, y.

(c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

This point is consideren an outlier because its value is very far from the predicted value of the least-squares regression line. The point's value is about 100 while, when there are 3 customers in line, the LSRL predicts a value of about 600.

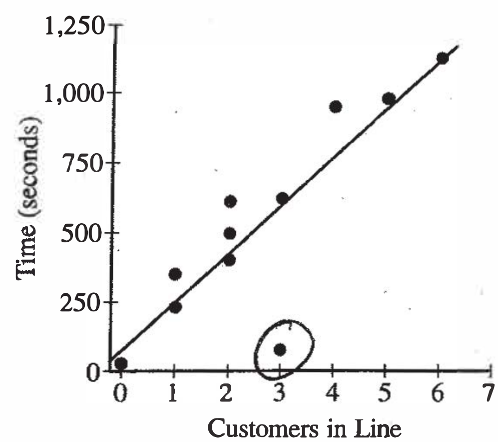
GO ON TO THE NEXT PAGE.

STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 1 hour and 5 minutes on this part of the exam.

Percent of Section II score—75

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.

1. The manager of a grocery store selected a random sample of 11 customers to investigate the relationship between the number of customers in a checkout line and the time to finish checkout. As soon as the selected customer entered the end of a checkout line, data were collected on the number of customers in line who were in front of the selected customer and the time, in seconds, until the selected customer was finished with the checkout. The data are shown in the following scatterplot along with the corresponding least-squares regression line and computer output.



Predictor	Coef	SE Coef	T	P
Constant	72.95	110.36	0.66	0.525
Customers in line	174.40	35.06	4.97	0.001
S = 200.01		R-Sq = 73.33%	R-Sq (adj) = 70.37%	

GO ON TO THE NEXT PAGE.

1C 2

(a) Identify and interpret in context the estimate of the intercept for the least-squares regression line.

The estimate given for the intercept is 72.95 seconds. This means that with 0 customers in line, one would expect the time to finish checkout to be approximately 72.95 seconds.

(b) Identify and interpret in context the coefficient of determination, r^2.

r^2 is given by the output as 70.37%, or 0.7037. This indicates that approximately 70.37% of the variance of results from the expected times (regression line) is accounted for by the least-squares regression line.

(c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

The point at approximately (3, 100) is considered an outlier because it is significantly further from the regression line, or expected time for its number of customers, than any other point. This negatively affects the accuracy of the regression line.

GO ON TO THE NEXT PAGE.

Question 1

Overview

The primary goals of this question were to assess a student's ability to (1) identify various values in regression computer output; (2) interpret the intercept of a regression line in context; (3) interpret the coefficient of determination in context; and (4) identify an outlier from a scatterplot.

Sample: 1A
Score: 4

In part (a) the response correctly recognizes the value of the intercept, satisfying component 1. The response then communicates the concept of an intercept using context that incorporates both time and zero customers; this satisfies component 2. Because the interpretation of the intercept indicates that the value is a prediction, as indicated by "we would expect them to be finished with checkout in 72.95 seconds," component 3 is satisfied. This response includes all three components; therefore, part (a) was scored as essentially correct. In part (b) the response correctly recognizes the value of r^2, satisfying component 1. The response provides a correct interpretation of r^2; this satisfies component 2. Because the interpretation is made using context, component 3 is satisfied. The response includes all three components; therefore, part (b) was scored as essentially correct. In part (c) the outlier is circled on the scatterplot, satisfying component 1. The response gives valid reasoning why the circled point is the outlier, relative to the remaining data points, by stating the circled point "does not follow the pattern the rest of the data points follow," and that the circled point is "very far from the rest of the data." Component 2 is satisfied. The response includes both components; therefore, part (c) was scored as essentially correct. Because three parts were scored as essentially correct, the response earned a score of 4.

Sample: 1B
Score: 3

In part (a) the response correctly recognizes the value of the intercept, satisfying component 1. The response also communicates the concept of an intercept using context that incorporates both time and zero customers; this satisfies component 2. Because the interpretation of the intercept indicates that the value is a prediction, as indicated by "the predicted time to finish checkout," component 3 is satisfied. The response includes all three components; therefore, part (a) was scored as essentially correct. In part (b) the response correctly recognizes the value of r^2, satisfying component 1. The response provides a correct interpretation of r^2; this satisfies component 2. Because the interpretation is made using context, component 3 is satisfied. The response includes all three components; therefore, part (b) was scored as essentially correct. In part (c) the outlier is circled on the scatterplot, satisfying component 1. The response does not, however, give sufficient reasoning to explain why the circled point is the outlier. The statement, "its value is very far from the predicted value of the least-squares regression line," does not make a comparison to the variation in the remaining data points, nor does it reference the distance between the y-coordinate of the circled point and the value of the prediction at $x = 3$. Therefore, in the absence of any comparison against the distances of all other points to the regression line, the response does not satisfy component 2. The response includes component 1 but not component 2; therefore, part (c) was scored as partially correct. Because two parts were scored as essentially correct, and one part was scored as partially correct, the response earned a score of 3.

Question 1 (continued)

Sample: 1C
Score: 2

In part (a) the response correctly recognizes the value of the intercept, satisfying component 1. The response then communicates the concept of an intercept using context that incorporates both time and zero customers; this satisfies component 2. The interpretation of the intercept indicates that the value is a prediction in two different ways; the first is indicated by "one would expect the time to finish checkout to be," and the other is indicated by "approximately." Either of these indications satisfies component 3. The response includes all three components; therefore, part (a) was scored as essentially correct. In part (b) the response gives an incorrect value of r^2, so component 1 is not satisfied. The response incorrectly interprets r^2 in terms of the percent of the variance in "results from the expected times," instead of percent variance in the observed times. Therefore the response does not satisfy component 2. Because the interpretation is made using context, component 3 is satisfied. The response includes only one of the three components, consequently, part (b) was scored as incorrect. In part (c) the outlier is circled on the scatterplot, satisfying component 1. The response gives valid reasoning why the circled point is the outlier by stating the circled point is "significantly farther from the regression line … than any other point." It is this portion of the response that compares the circled point to the remaining data points and, therefore, the response satisfies component 2. Because the response includes both components, part (c) was scored as essentially correct. Because two parts were scored as essentially correct, and one part was scored as incorrect, the response earned a score of 2.

2018

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free Response Question 2

☑ Scoring Guideline

☑ Student Samples

☑ Scoring Commentary

Question 2

Intent of Question

The primary goals of this question were to assess a student's ability to (1) calculate the sample size when given the endpoints of a confidence interval for a proportion; (2) explain how bias could be present in a particular survey method; and (3) estimate a proportion from sample data collected using a method designed to decrease bias.

Solution

Part (a):

Using the standard formula for a confidence interval for one proportion, the interval (0.584 to 0.816) is found as follows. $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ where $\hat{p} = \frac{0.584 + 0.816}{2} = 0.7,$ the margin of error is $0.816 - 0.7 = 0.116,$ and $z^* = 1.96.$

Solving $1.96\sqrt{\frac{0.7(1-0.7)}{n}} = 0.116$ yields $n = \frac{(1.96)^2(0.7)(1-0.7)}{(0.116)^2} \approx 59.95.$ The sample size was 60.

Part (b):

Bias might have been introduced because students responded directly to the environmental science teacher. Because the students would know that an environmental science teacher cares about the environment, they might say yes when they actually don't recycle. This would result in a point estimate that is greater than the proportion of all students who would respond yes to the question.

Part (c):

(i) The expected number is $(300)\left(\frac{1}{2}\right) = 150.$

(ii) The point estimate is based on expecting 150 students to be required to say no and 150 students to truthfully answer the question. Of the 213 answers of no, we expect that $213 - 150 = 63$ were from students who truthfully answered the question. That means we expect that the remaining $150 - 63 = 87$ students truthfully answered the question and responded yes. So the point estimate for the proportion of all students at the high school who would respond yes to the question is $\frac{87}{150} = 0.58.$

Question 2 (continued)

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following five components:

1. Uses a standard error in the form $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ where $\hat{p}$ is between 0 and 1.
2. Shows evidence that $\hat{p} = 0.7$ was correctly used in the standard error.
3. Shows evidence that 0.116 was correctly used as the margin of error in the calculation.
4. Shows evidence that $z^* = 1.96$ was correctly used as the critical value in the calculation.
5. Includes a single, positive whole-number answer.

Partially correct (P) if the response satisfies only three or four of the five components.

Incorrect (I) if the response satisfies at most two of the five components.

Notes:

- Using an equation in the form $n = \frac{z^2\hat{p}(1-\hat{p})}{\text{MOE}^2}$ satisfies component 1.
- A value of 0.21 in the numerator of the standard error implies that $\hat{p} = 0.7$ was correctly used in the standard error and satisfies component 2.
- An equation such as $0.816 = 0.7 + \text{MOE}$ implies that 0.116 was correctly used for the margin of error and satisfies component 3.
- Statements that suggest a whole-number answer is approximate (such as, “about 60” or “$\approx$ 60”) satisfy component 5.
- Algebraic work between the set-up and final answer does not need to be shown to satisfy component 5.
- When calculating the values 0.7, 0.116, or 1.96, ignore minor arithmetic errors or transcription errors if they can be identified by the work shown.

Question 2 (continued)

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Explains why the responses to the survey might differ from the truth about student recycling in this context (for example, the survey was not anonymous, the question was asked by an authority figure).
2. Explains how the responses to the survey might differ from the truth about student recycling (for example, "students might say yes when they actually don't recycle," "students lie and say yes," "students don't recycle but lie to the teacher").
3. Describes the effect of the bias on the point estimate (or the proportion, percentage, number of yes responses in the sample) and doesn't contradict the bias described.

Partially correct (P) if the response satisfies only two of the three components.

Incorrect (I) if the response satisfies at most one of the three components.

Notes:

- To satisfy component 1 the response must provide a reason that is based on a bias created by the teacher asking students in person. For example, a response that addresses the wording of the question, voluntary response, or sampling variability does not satisfy component 1.
- To satisfy component 2 the response needs to explicitly contrast what the students say with what they do.
- Evidence used to address component 3 cannot also be used to address component 2. For example, a response that says "Students might lie, producing an estimate that is too high" addresses the effect of the bias on the point estimate but should not be combined with the statement about students lying to infer that students do not actually recycle. However, a response that says "Students may lie and say yes, producing an estimate that is too high" satisfies both components 2 and 3.
- If the response is clearly about the population proportion and not about the point estimate, component 3 cannot be satisfied.
- Statements such as "the interval will be too high" do not satisfy component 3 because they don't specifically address the point estimate.

Question 2 (continued)

Part (c) is scored as follows:

Essentially correct (E) if the response gives an answer of 150 in (c-i) and gives an answer of 0.58 (or equivalent) in (c-ii).

Partially correct (P) if the response gives an answer of 150 in (c-i) and gives an answer of 0.42 (or equivalent) in (c-ii);

OR

if the response does not give an answer of 150 in (c-i) but gives an answer of 0.58 (or equivalent) with supporting work in (c-ii).

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- In part (c-i) the answer must be a single number. Responses such as "at least 150" or "147–153" are incorrect. However, responses such as "about 150" or "$\approx$ 150" are acceptable.
- In part (c-ii) the proportion can be described verbally (e.g., "87 out of 150").
- In part (c-ii) if the response clearly indicates that 0.58 (or 0.42) is the population proportion, lower the overall score in part (c) by one level (that is, from E to P, or from P to I). Using probability notation such as P(yes) does not clearly indicate a population proportion.
- In part (c-ii) if the response includes a point estimate of 0.58 or 0.42 but uses a confidence interval as the final answer, lower the overall score in part (c) by one level (that is, from E to P, or from P to I).
- If the answer is incorrect in part (c-i) and the answer in part (c-ii) uses numerator = 87 and denominator = 300 – answer to (c-i), the response should be scored P.

Question 2 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and one or two parts partially correct

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?

$0.816 - 0.584 = 0.232$

$0.232 \div 2 = 0.116 \leftarrow$ margin of error

$0.584 + 0.116 = 0.7 = p$

$1 - 0.7 = 0.3 = q$

$\frac{1-0.95}{2} = 0.025$

invNorm

area: 0.025
μ: 0
σ: 1

$\Big\}\ 1.96 = z^*$

$0.7 \pm z^* \sqrt{\frac{pq}{n}}$

$0.7 \pm 1.96 \sqrt{\frac{(0.7)(0.3)}{n}}$

margin of error $\rightarrow 0.116 = 1.96 \sqrt{\frac{(0.7)(0.3)}{n}}$

$n \approx 60$ students

60 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.

Bias might have been introduced in terms of response bias. The environmental science teacher is likely to support regularly recycling plastic bottles, so students might have been more likely to respond yes to the question, even if that was not their true answer. This bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question by causing the point estimate to be too high. Due to the pressure to answer a certain way (yes) that the students in the sample likely felt, it is likely that the point estimate was too high.

GO ON TO THE NEXT PAGE.

(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?

$300\left(\frac{1}{2}\right) = 150$

The expected number of students from the sample of 300 who would be required to respond no because the coin flip resulted in heads is 150 students. Since there are only 2 potential outcomes to the coin flip, about half of the sample of 300 would be expected to get a coin flip of heads and be required to respond no.

(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.

$213 - 150 = 63$

About 63 students got a coin flip of tails, meaning they truthfully responded no.

$150 - 63 = 87$

$\frac{87}{150} = 0.58$

Out of the 150 students who got a coin flip of tails, about 63 of them responded no. Thus, the other 87 students responded yes truthfully (a proportion of 0.58).

A point estimate for the proportion of all students at the high school who would respond yes to the question is 0.58

2B1 2B1

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?

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

$\hat{p} = \frac{.584 + .816}{2} = 0.7$

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

$MOE = .816 - .7 = .116$

$.116 = 1.96 \sqrt{\frac{.7(.3)}{n}}$

$\frac{.116}{1.96} = \sqrt{\frac{.7(.3)}{n}}$

$\frac{.7(.3)}{n} = \left(\frac{.116}{1.96}\right)^2$

$n = \left(\frac{1.96}{.116}\right)^2 (.7)(.3) = 59.95$

≈ 60 students

(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.

The students might feel pressure to answer yes since they were being surveyed by an environmental science teacher who would likely promote recycling. Thus, the point-estimate might overestimate the true proportion of all students who recycle plastic bottles.

2B2 2B2

(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?

$(.5)(300) = 150$ students

P(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.

expected number of "no" responses from flipping heads

$213 - 150 = 63$

we expect 63 students to have responded no after flipping tails, and ~~these~~ would be the truthful responses

$\hat{p}_{\text{honest no}} = \frac{63}{150} = .42$

$\hat{p}_{yes} = 1 - .42 = \boxed{.58}$

GO ON TO THE NEXT PAGE.

2C1 2C1

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?

point estimate = $(0.816 + 0.584)/2 = 0.7$

$$(0.584, 0.816) = 0.7 \pm 1.96\sqrt{\frac{(0.7)(0.3)}{n}}$$

$$0.584 = 0.7 - 1.96\sqrt{\frac{(0.7)(0.3)}{n}}$$

$$\boxed{n \approx 60 \text{ students}}$$

(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.

Since the teacher asking the question is an environmental science teacher, the students may feel that they need to show that they recycle regularly when in reality they might not. Due to this bias, the point estimate of the proportion may be higher than it actually is. This is an example of a response bias.

GO ON TO THE NEXT PAGE.

2C2

(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?

$300(0.5) = \boxed{150 \text{ students}}$

(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.

$\frac{(300-213)}{300} = \boxed{0.29}$

GO ON TO THE NEXT PAGE.

Question 2

Overview

The primary goals of this question were to assess a student's ability to (1) calculate the sample size when given the endpoints of a confidence interval for a proportion; (2) explain how bias could be present in a particular survey method; and (3) estimate a proportion from sample data collected using a method designed to decrease bias.

Sample: 2A
Score: 4

In part (a) the response includes the formula $\sqrt{\frac{pq}{n}}$ for the standard error, satisfying component 1. Because the difference between a point estimate and population proportion is assessed in parts (b) and (c), using p and q instead of $\hat{p}$ and $\hat{q}$ in the formula was considered a minor error. In the equation $0.116 = 1.96\sqrt{\frac{(0.7)(0.3)}{n}}$, the response correctly uses $\hat{p} = 0.7$ in the standard error, which satisfies component 2, correctly uses 0.116 as the margin of error, which satisfies component 3, and correctly uses 1.96 as the critical value, which satisfies component 4. Finally, the response includes a single, positive whole-number answer (60), satisfying component 5. Because the response includes all five components, part (a) was scored as essentially correct. In part (b) the response provides a reason why the survey responses might be different from the truth ("Due to the pressure to answer a certain way"), satisfying component 1. The response explains how the survey responses differ from the truth ("more likely to respond yes to the question, even if that was not their true answer"), satisfying component 2. Finally, the response describes the effect of the bias on the point estimate ("causing the point estimate to be too high"), satisfying component 3. Because the response includes all three components, part (b) was scored as essentially correct. In part (c-i) the response gives an answer of 150. In part (c-ii) the response provides an answer of 0.58. Because the response includes both correct answers, part (c) was scored as essentially correct. Because three parts were scored as essentially correct, the response earned a score of 4.

Sample: 2B
Score: 3

In part (a) the response includes the equation $0.116 = 1.96\sqrt{\frac{.7(.3)}{n}}$, which uses a standard error in the correct form, satisfying component 1. In the same equation, the response also correctly uses $\hat{p} = 0.7$ in the standard error, which satisfies component 2, correctly uses 0.116 as the margin of error, which satisfies component 3, and correctly uses 1.96 as the critical value, satisfying component 4. Finally, the response includes a single, positive whole-number answer (60), satisfying component 5. Because the response includes all five components, part (a) was scored as essentially correct. In part (b) the response provides a reason why the survey responses might be different than the truth ("the students might feel pressure to answer yes"), satisfying component 1. While the response indicates that the students are likely to say yes ("pressure to answer yes"), the response does not indicate that the students do not actually recycle. Because the response does not describe how the survey responses would differ from the truth, component 2 is not satisfied. Finally, the response describes the effect of the bias on the point estimate ("the point-estimate might overestimate the true proportion of all students who recycle plastic bottles"), satisfying component 3. Because the response includes two of the three components, part (b) was scored as partially correct. In part (c-i) the response gives an answer of 150. In part (c-ii) the response provides an answer of 0.58. Because the response includes both correct answers, part (c) was scored as essentially correct. Because two parts were scored as essentially correct, and one part was scored as partially correct, the response earned a score of 3.

Question 2 (continued)

Sample: 2C
Score: 2

In part (a), the response includes the formula $0.584 = 0.7 - 1.96\sqrt{\frac{(0.7)(0.3)}{n}}$, which uses a standard error in the correct form, satisfying component 1. In the same equation, the response also correctly uses $\hat{p} = 0.7$ in the standard error, which satisfies component 2, and correctly uses 1.96 as the critical value, which satisfies component 4. Furthermore, the equation implies that 0.116 was correctly used for the margin of error (described in the third scoring note of the scoring guidelines), satisfying component 3. Finally, the response includes a single, positive whole-number answer (60), satisfying component 5. Because the response includes all five components, part (a) was scored as essentially correct. In part (b) the response provides a reason why the survey responses might be different from the truth ("Since the teacher asking the question is an environmental science teacher"), satisfying component 1. The response explains how the survey responses differ from the truth ("show that they recycle regularly when in reality they might not"), satisfying component 2. Finally, the response describes the effect of the bias on the point estimate ("the point estimate of the proportion may be higher than it actually is"), satisfying component 3. Because the response includes all three components, part (b) was scored as essentially correct. In part (c-i) the response gives an answer of 150. In part (c-ii) the response provides an answer of 0.29, which is the overall proportion of "yes" responses in the sample. Because the response in part (c-ii) does not account for the 150 students who would be expected to say "no" because of the coin flip, the response was scored as incorrect. Because two parts were scored as essentially correct, and one part was scored as incorrect, the response earned a score of 2.

2018

AP CollegeBoard

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free Response Question 3

☑ Scoring Guideline

☑ Student Samples

☑ Scoring Commentary

Question 3

Intent of Question

The primary goals of this question were to assess a student's ability to (1) compute a probability based on a weighted mixture of two populations; (2) compute a conditional probability; and (3) recognize a binomial random variable and compute the probability associated with it.

Solution

Part (a):

Let L denote left-handed, M denote multiple birth, and S denote single birth.
The probability that a randomly selected child born in the region is left-handed is:

$$P(L) = P(M)P(L \mid M) + P(S)P(L \mid S) = (0.035)(0.22) + (0.965)(0.11) = 0.0077 + 0.10615 = 0.11385.$$

Part (b):

From part (a), $P(L) = 0.11385.$ Therefore,

$$P(M \mid L) = \frac{P(L \text{ and } M)}{P(L)} = \frac{(0.035)(0.22)}{0.11385} = \frac{0.0077}{0.11385} \approx 0.0676.$$

Part (c):

Let X represent the number of children who are left-handed in a random sample of 20 children from the region. X has a binomial distribution with $n = 20$ and $p = 0.11385$ (found in part (a)). Using the binomial distribution,

$$P(X \geq 3) = 1 - P(X \leq 2)$$
$$= 1 - \binom{20}{0}(0.11385)^0(0.88615)^{20} - \binom{20}{1}(0.11385)^1(0.88615)^{19} - \binom{20}{2}(0.11385)^2(0.88615)^{18}$$
$$\approx 1 - 0.598 \approx 0.402.$$

Question 3 (continued)

Scoring

Parts (a), (b), and (c) each scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the probability is computed correctly, *AND* work is shown that includes correct numerical values using a formula, end results from a tree diagram, or some other appropriate strategy.

Partially correct (P) if the response provides a reasonable strategy for finding the probability, such as a formula or tree diagram, but uses one or more inappropriate values;

OR

if the response gives the correct probability but not enough work is shown to determine how that probability was found.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: A reasonable strategy needs to include summing the results of two multiplications.

Part (b) is scored as follows:

Essentially correct (E) if the probability is computed correctly, with work shown that includes appropriate numerical values for both the numerator and denominator.

Partially correct (P) if the response includes a numerator and denominator in calculating the conditional probability, with one appropriate term (numerator or denominator) and the other inappropriate.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: Appropriate values include incorrectly calculated values from part (a).

Question 3 (continued)

Part (c) is scored as follows:

Essentially correct (E) if the response satisfies the following five components:
1. Uses a calculation based on the binomial distribution to find the probability of the number of children in the sample who are left-handed.
2. Specifies appropriate values for n and p.
3. Uses correct endpoint value for the probability.
4. Uses correct direction to calculate the probability of at least three left-handed children.
5. Correctly calculates a binomial probability consistent with the previous work.

Partially correct (P) if the response satisfies component 1 and only two or three of the other four components;
OR
if components 2, 3, 4, and 5 are met, and the response does not explicitly indicate the binomial distribution is used by name or formula.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:
- "Appropriate" values include incorrectly calculated values from part (a) or a recalculated probability from part (b).
- An unlabeled numerical value in a calculator statement cannot be used to satisfy a component.
- A response which calculates $P(X \leq 3)$ satisfies component 3 but does not satisfy component 4.
- A normal approximation to the binomial is not appropriate because $np = 20 \times 0.11385 = 2.277 < 5$. A response using the normal approximation can score at most P. To earn a score of P, the response must include all of the following:
    - a correct mean and standard deviation based on the binomial parameters
    - clear indication of boundary and direction with a z-score or diagram
    - the probability computed correctly

Notes for all parts:
- If the resulting probability or part of the calculation of the probability uses a value that is not between 0 and 1, inclusive, the score for that part is lowered by one level (that is, from E to P, or from P to I).
- An arithmetic or transcription error in a response can be ignored if correct work is shown. For example, $0.0077 + 0.10615 = 0.1385$.

Question 3 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

3A1

3. Approximately 3.5 percent of all children born in a certain region are from multiple births (that is, twins, triplets, etc.). Of the children born in the region who are from multiple births, 22 percent are left-handed. Of the children born in the region who are from single births, 11 percent are left-handed.

(a) What is the probability that a randomly selected child born in the region is left-handed?

Let M = multiple   L = left
S = single   R = right

.035 multiple birth
  .22 left-handed $P(M \cap L) = (.035)(.22) = .0077$
  .78 right-handed $P(M \cap R) = (.035)(.78) = .0273$
.965 single birth
  .11 left-handed $P(S \cap L) = (.965)(.11) = .10615$
  .89 right handed $P(S \cap R) = (.965)(.89) = .85885$

$P(L) = .11385$

$P(R) = .88615$

$P(\text{left handed}) = [(.035)(.22)] + [(.965)(.11)]$

$\boxed{P(\text{left-handed}) = .1139}$

(b) What is the probability that a randomly selected child born in the region is a child from a multiple birth, given that the child selected is left-handed?

* refer to tree above

$$P(M|L) = \frac{P(M \cap L)}{P(L)}$$

$$P(M|L) = \frac{.0077}{.1139}$$

$\boxed{P(M|L) = .0676}$

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

3A2

(c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

$n=20 \quad p(L) = .1139$

$p(X\geq 3) = 1 - [(p(X=0)) + p(X=1) + p(X=2)]$

$p(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

$p(X=0) = \binom{20}{0}(.1139)^0 (.8861)^{20} = .0891$

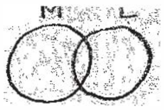
$p(X=1) = \binom{20}{1}(.1139)^1 (.8861)^{19} = .2289$

$p(X=2) = \binom{20}{2}(.1139)^2 (.8861)^{18} = .2796$

$p(X\geq 3) = 1 - (.0891 + .2289 + .2796)$

$\boxed{p(X\geq 3) = .4024}$

GO ON TO THE NEXT PAGE.



$P(M \cap L) = .22$ $P(S \cap L) = .11$

3B1

.035 M

P(M)

3. Approximately 3.5 percent of all children born in a certain region are from multiple births (that is, twins, triplets, etc.). Of the children born in the region who are from multiple births, 22 percent are left-handed. Of the children born in the region who are from single births, 11 percent are left-handed.

P(L)

(a) What is the probability that a randomly selected child born in the region is left-handed?

$P(L) = .22 + .11 = .33$

The probability a randomly selected child is left handed is .33.

(b) What is the probability that a randomly selected child born in the region is a child from a multiple birth, given that the child selected is left-handed?

$$P(M|L) = \frac{P(M \cap L)}{P(L)} = \frac{.22}{.33} = .6667$$

The probability a child is left handed given they're from a multiple birth is .6667

GO ON TO THE NEXT PAGE.

3B2

(c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

trials prob X

$P(L \geq 3) = 1 - P(L < 2) = \text{binomcdf}(20, .067, 2)$

2.26×10^{-7}

The probability that out of 20 children in the region at least three are left-handed is 2.26×10^{-7}

GO ON TO THE NEXT PAGE.

3C1

3. Approximately 3.5 percent of all children born in a certain region are from multiple births (that is, twins, triplets, etc.). Of the children born in the region who are from multiple births, 22 percent are left-handed. Of the children born in the region who are from single births, 11 percent are left-handed.

(a) What is the probability that a randomly selected child born in the region is left-handed?

$\hat{p}_A = .22$

$\hat{p}_B = .11$ $\quad 0.22 + 0.11 = 0.33$

The probability that a randomly selected child born in this region is left handed is 0.33 or 33%.

(b) What is the probability that a randomly selected child born in the region is a child from a multiple birth, given that the child selected is left-handed?

$P(A|B)$ $\quad P(A) = .035$, $P(B) = .33$ $\quad P(A \cap B) = .035(0.33) = .012$

Where P(A) is the probability that the child is from a multiple birth and P(B) is the probability that the child is left handed.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.012}{0.33} = .036$$

The probability that a randomly selected child born in the region is a child from a multiple birth given that the child selected is left handed is .036 or 3.6%.

GO ON TO THE NEXT PAGE.

3C2

(c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

$\binom{20}{3}$ 20 nCr 3

$1140(.33)^3(1-.33)^{17} = .045$

The probability that a random sample of 20 children born in the region will result in at least 3 children who are left handed is .045 or 4.5%.

GO ON TO THE NEXT PAGE.

Question 3

Overview

The primary goals of this question were to assess a student's ability to (1) compute a probability based on a weighted mixture of two populations; (2) compute a conditional probability; and (3) recognize a binomial random variable and compute the probability associated with it.

Sample: 3A
Score: 4

In part (a) the response has a well-labeled tree diagram with the events and corresponding probabilities on the branches. The multiplication is shown at the end of each branch using correct notation, and the symbols are defined above the tree. The response clearly indicates which branches have been used to find $P(L)$ at the end of the tree and then summarizes the arithmetic below the tree. Part (a) was scored as essentially correct. In part (b) the same symbols used in part (a) are found. The formula for the conditional probability is given using the symbols defined in part (a). The probabilities, as found in the tree and corresponding to the symbols, are given in the numerator and denominator of the fraction. The final probability is correctly calculated. Part (b) was scored as essentially correct. In part (c) the response indicates how the probability of x equal to k is found. This statement satisfies component 1. The response clearly indicates the value for n and the probability of being left-handed as found in part (a). These statements satisfy component 2. The response gives the statement $p(x \geq 3) = 1 - [(p(x = 0)) + p(x = 1) + p(x = 2)]$, which includes the correct value for the endpoint of the cumulative distribution and the direction used to calculate the probability, satisfying components 3 and 4, respectively. Although not necessary, each of these calculations is shown. The final probability is correctly calculated, satisfying component 5. Part (c) was scored as essentially correct. Because the three parts were scored as essentially correct, the response earned a score of 4.

Sample: 3B
Score: 2

In part (a) the response does not weight the conditional probability of left-handed children of multiple births by the probability that a child is from a multiple birth nor does the response weight the conditional probability of the left-handed children of single births by the probability the child is from a single birth. Because this is not a reasonable strategy to find a weighted probability, part (a) was scored as incorrect. In part (b) the correct formula for a conditional probability is given. The incorrect value for $P(M \cap L)$ from part (a) is used and is identified at the top of the page. The incorrect value for $P(L)$ from part (a) is used. Because the response clearly identifies the quantities and carries the quantities from part (a), the quantities were scored as appropriate as specified in the note for part (b) in the scoring guidelines. The statement "The probability a child is left handed given they're from a multiple birth" was overlooked in the scoring. Part (b) was scored as essentially correct. In part (c) the response uses the calculator notation for the function binomcdf(20, .667, 2) and labels the parameters. Component 1 is satisfied with the function statement. Labeling 20 as trials satisfies the n portion of component 2. However, the value labeled as "prob" is the value labeled in part (b) as "probability a child is left handed given they're from a multiple birth." In part (a) the response defines the "probability a randomly selected child is left-handed is .33." In part (c) the response specifies binomcdf(20,.667,2) and labels the parameters, satisfying component 1. The probabilities defined in parts (a) and (b) are used and component 2 is satisfied. Component 3 is satisfied with the statement $P(L \geq 3)$. There is poor notation in the equation as $1 - P(L < 2)$ should equal $1 - (\text{binomcdf}(20,.667,2)$, but that was overlooked. The calculator notation does indicate the correct endpoint for the complement and the correct direction. Component 4 is satisfied. The calculation is incorrect for the given p, so component 5 is not satisfied. Because component 1 and two of the other four components are satisfied, part (c) was scored as partially correct. Because one part was scored as essentially correct, one part was scored as partially correct, and one part was scored as incorrect, the response earned a score of 2.

Sample: 3C
Score: 1

In part (a) the response does not weight the conditional probability of left-handed children of multiple births by the probability that a child is from a multiple birth nor does the response weight the conditional probability of the left-handed children of single births by the probability the child is from a single birth. Because this is not a reasonable strategy to find a weighted probability, part (a) was scored as incorrect. In part (b) the probability $P(B)$ calculated in part (a) is used. The probability that a child is left-handed and from a multiple birth is incorrect because these events are not independent. Because the numerator of the conditional probability is incorrect, part (b) was scored as partially correct. In part (c) the binomial distribution is attempted with the formula given, satisfying component 1. Because the combination $\binom{20}{3}$ is standard notation, component 3 is satisfied, and the portion of component 2 requiring the identification of n is satisfied. The portion of component 2 requiring p is satisfied because the probability from part (a) is used in the binomial formula. The probability of $x = 3$ is calculated correctly, satisfying component 5. However, no direction is given, and component 4 is not satisfied. Because component 1 and three of the other four components are satisfied, part (c) was scored as partially correct. Because two parts were scored as partially correct, and one part was scored as incorrect, the response earned a score of 1.

2018

AP CollegeBoard

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free Response Question 4

- ☑ Scoring Guideline
- ☑ Student Samples
- ☑ Scoring Commentary

Question 4

Intent of Question

The primary goals of this question were to assess a student's ability to (1) determine whether a cause-and-effect conclusion can be made based on how a study was conducted and (2) set up, perform, and interpret the results of a hypothesis test, in the context of the problem.

Solution

Part (a):

Yes, it would be reasonable to conclude that the new procedure causes a reduction in recovery time, for patients similar to those in the study. The patients in the study were randomly assigned to the two procedures, which reduces the chance that confounding variables will affect the results. Therefore the statistically significant reduction in mean recovery time can be attributed to the new procedure being superior to the standard procedure.

Part (b):

Step 1: State a correct pair of hypotheses.

Let μ_S represent the mean recovery time among all patients similar to those in the study if they were to receive the standard treatment.

Let μ_N represent the mean recovery time among all patients similar to those in the study if they were to receive the new treatment.

The hypotheses to be tested are $\mathrm{H}_0 : \mu_S = \mu_N$ versus $\mathrm{H}_a : \mu_S > \mu_N$.

Step 2: Identify a correct test procedure (by name or by formula) and check appropriate conditions.

The appropriate procedure is a two-sample *t*-test for a difference between means.

Because this is an experiment, the first condition is that subjects were randomly assigned to one treatment group or the other. In this case the condition is satisfied because we were told that the subjects were randomly assigned to either the standard or new procedure.

The second condition is that the recovery times of the two populations are normally distributed or the sample sizes are sufficiently large to presume that the distribution of the difference in the sample means is approximately normal. In this case the condition is met because the sample sizes of 110 and 100 are both sufficiently large.

Question 4 (continued)

Step 3: Correct mechanics, including the value of the test statistic, degrees of freedom, and p-value (or rejection region).

The test statistic is $t = \dfrac{\bar{x}_S - \bar{x}_N}{\sqrt{\dfrac{s_S^2}{n_S} + \dfrac{s_N^2}{n_N}}} = \dfrac{217 - 186}{\sqrt{\dfrac{34^2}{110} + \dfrac{29^2}{100}}} \approx 7.13$.

The p-value is the area greater than 7.13 for a t-distribution with $df = 207.18$, which is essentially 0 $\left(8.36 \times 10^{-12}\right)$.

Step 4: State a correct conclusion in the context of the problem, using the result of the statistical test.

Because the p-value is very small, we have sufficient evidence to conclude that for patients similar to the ones in the study, those receiving the new procedure would have less recovery time, on average, than those receiving the standard procedure.

Scoring

This question is scored in three sections. Section 1 consists of part (a); section 2 consists of step 1, step 2, and the test statistic in step 3 in part (b); and section 3 consists of the p-value in step 3 and step 4 in part (b). Sections 1, 2, and 3 are each scored essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Correctly states that it is reasonable to make a causal conclusion.
2. Justifies the causal conclusion based on random assignment of patients to procedures (or procedures to patients);
   OR
   justifies the causal conclusion by stating that a randomized experiment was conducted.
3. Includes the context of the situation.

Partially correct (P) if the response satisfies component 1 *AND* provides WEAK justification of the causal conclusion by stating that there was random assignment or a randomized experiment was conducted, but with no context;

OR

by stating that an experiment was conducted or there was assignment (without the word "randomized") *AND* the response includes context of the situation;

OR

by stating that the study design reduces the chance of confounding variables or balances the effects of uncontrolled variables across both groups in context without explicitly referring to the random assignment.

Incorrect (I) if the response does not meet the criteria for E or P.

Question 4 (continued)

Notes:
- If the response states that it is *not* reasonable to make a causal conclusion because the result could have been due to random chance *AND* explains that there is evidence for a causal conclusion based on random assignment of patients to procedures or by stating that a randomized experiment was conducted, then the response is scored E.
- If the response discusses aspects of an experiment other than random assignment (such as, control, replication, or large samples), then those aspects are considered extraneous and the response can be scored E unless those aspects are incorrect for this study (such as, blocking is a requirement, or the study used blocking, or the study used a placebo) in which case the score should be lowered one level (that is, from E to P, or from P to I).
- If the response correctly states in context that it is reasonable to make a causal conclusion but includes incorrect or contradictory justification (such as, random selection of patients), then the response is scored I.

Section 2 is scored as follows:

Essentially correct (E) if the response satisfies the following four components:
1. Parameters are defined correctly.
2. Hypotheses imply equality in the null and correct direction in the alternative.
3. Correct test is identified by name or formula.
4. Correct test statistic for a difference in means is calculated.

Partially correct (P) if the response satisfies only two or three of the four components.

Incorrect (I) if the response satisfies at most one of the four components.

Notes:
- If standard symbols are used for the parameters with appropriate group labels (such as, μ_S, μ_N), component 1 is satisfied.
- If the correct test is identified, but the response states an incorrect formula or uses incorrect notation in the formula, component 3 is not satisfied.
- A pooled two-sample *t*-test is acceptable for component 3, but the student must also state and comment on the plausibility of the equal population variances assumption.
- If the response identifies a *z*-test for equal means as the correct test identification, component 3 is not satisfied but component 4 could be satisfied.

Confidence Interval approach:
- If a single two-sample *t*-interval for the difference in means is used, components 3 and 4 can be satisfied. Component 3 is satisfied if the *t*-interval is correctly identified by name or formula. Component 4 is satisfied if the correct interval is calculated. If an alpha level is stated, then an appropriate adjustment to the confidence level must be made because the appropriate test is one-sided.
- If two one-sample *t*-intervals are used, while not a recommended approach, component 3 is not satisfied but component 4 could be satisfied. Component 4 is satisfied if both intervals are calculated correctly.

Question 4 (continued)

Section 3 is scored as follows:

Essentially correct (E) if the response satisfies the following three components:
1. Makes reference to an approximately correct *p*-value that is consistent with the test statistic and alternative hypothesis for a difference in means.
2. Correctly justifies the conclusion based on the size of the *p*-value or the test statistic.
3. Correctly states the conclusion in context.

Partially correct (P) if the response satisfies only two of the three components.

Incorrect (I) if the response does not meet the criteria for E or P or includes a justification not based on the inferential results.

Notes:
Component 1:
- Is satisfied if the response makes reference to a large test statistic without referring to a *p*-value.

Component 2:
- No alpha level is needed to provide justification of the conclusion based on the size of the *p*-value.
- Is satisfied if the response states the *p*-value without reference to size, but it is contiguous to the conclusion and clearly indicates a continuous train of thought.
- A correct interpretation of the *p*-value with a complete explanation that obtaining a test statistic at least this extreme is unlikely due to chance alone is considered justification based on the size of the *p*-value.
- If an incorrect interpretation of the *p*-value is given, the score is lowered one level (that is, from E to P, or from P to I).
- A decision about the null hypothesis (reject H_0 or fail to reject H_0) is not required, but if an incorrect decision is stated based on the given *p*-value then component 2 is not satisfied.
- If a rejection region approach is used, a reasonable critical value replaces the *p*-value.

Component 3:
- A correct conclusion must be related to the alternative hypothesis in order to satisfy component 3.
- The following responses do not satisfy component 3:
  - States or implies that the null hypothesis is *accepted*
  - States or implies that the alternative hypothesis has been *proven*
  - States the conclusion in past tense (unless the response did not satisfy a component of section 2 for the use of past tense)

Question 4 (continued)

Confidence Interval approach:

- If a single two-sample *t*-interval for the difference in means is used:
  - Component 1 is satisfied if the response indicates that zero is either included or not included in the calculated interval.
  - Component 2 is satisfied if the response indicates that the bounds are either both above or both below zero (consistent with alternative hypothesis) and uses that as justification for the conclusion.
  - Component 3 is satisfied if the conclusion is stated in context.
- If two one-sample *t*-intervals are used (which is not recommended) the response is scored at most P if all three components are satisfied, otherwise scored I:
  - Component 1 is satisfied if the response states that the intervals do not overlap.
  - Component 2 is satisfied if the conclusion indicates that the confidence interval for the new procedure lies below the confidence interval for the standard procedure.
  - Component 3 is satisfied if the conclusion is stated in context.

Note: If the three sections of the response are scored as E, to earn a score of 4 as a complete response, both conditions in step 2 must be correctly stated and justified. Additional condition(s) inappropriate for a two-sample *t*-test must not be stated. Otherwise, the response earns a score of 3 a substantial response.

Question 4 (continued)

4 Complete Response

Three sections essentially correct with conditions for inference

3 Substantial Response

Three sections essentially correct without conditions for inference

OR

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 Minimal Response

One section essentially correct

OR

No sections essentially correct and one or two sections partially correct

4. The anterior cruciate ligament (ACL) is one of the ligaments that help stabilize the knee. Surgery is often recommended if the ACL is completely torn, and recovery time from the surgery can be lengthy. A medical center developed a new surgical procedure designed to reduce the average recovery time from the surgery. To test the effectiveness of the new procedure, a study was conducted in which 210 patients needing surgery to repair a torn ACL were randomly assigned to receive either the standard procedure or the new procedure.

(a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

Yes. Since the patients were randomly assigned to treatments in this experiment, a cause-and-effect relationship can be drawn and applied to patients similar to those in the study.

(b) Summary statistics on the recovery times from the surgery are shown in the table.

Type of Procedure	Sample Size	Mean Recovery Time (days)	Standard Deviation Recovery Time (days)
Standard	110	217	34
New	100	186	29

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

State: $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 > 0$

$\alpha = 0.05$

μ_1 is mean recovery time in days for the population of patients recieving the standard procedure, μ_2 is mean recovery time in days for population of patients recieving the new procedure

GO ON TO THE NEXT PAGE.

If you need more room for your work in part (b), use the space below.

Plan: If the conditions are met, we will use a 2-sample t-test for difference between means

Random - the treatments were randomly assigned

Normal/Large Sample - the sample size of both the standard and new procedure is greater than 30, so their sampling distributions are approximately Normal

Since this is a randomized experiment, the 10% condition doesn't need to be checked.

Do: $t = \dfrac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

$\bar{x}_1 = 217$ $s_1 = 34$ $n_1 = 110$
$\bar{x}_2 = 186$ $s_2 = 29$ $n_2 = 100$

$t = \dfrac{217 - 186}{\sqrt{\dfrac{34^2}{110} + \dfrac{29^2}{100}}} = 7.127$

$df = 100 - 1 = 99$

p-value = tcdf(lower = 7.127, upper = 1×10^{99}, df = 99)

p-value = $8.48 \times 10^{-11} \approx 0$

Conclude

Since the p-value of approximately 0 is less than the α of 0.05, we reject H_0 and have convincing evidence that patients who recieve the new procedure will have, on average, a shorter recovery time than those who recieve the standard procedure

GO ON TO THE NEXT PAGE.

4. The anterior cruciate ligament (ACL) is one of the ligaments that help stabilize the knee. Surgery is often recommended if the ACL is completely torn, and recovery time from the surgery can be lengthy. A medical center developed a new surgical procedure designed to reduce the average recovery time from the surgery. To test the effectiveness of the new procedure, a study was conducted in which 210 patients needing surgery to repair a torn ACL were randomly assigned to receive either the standard procedure or the new procedure.

(a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

Yes, because the patients were randomly assigned to treatment groups. Random assignment is a form of control that allows us to make inferences about cause and effect in a well-designed experiment because we hope it reduces any bias that is a result of confounding variables we didn't directly control, by creating groups that are roughly equivalent in terms of those variables.

(b) Summary statistics on the recovery times from the surgery are shown in the table.

Type of Procedure	Sample Size	Mean Recovery Time (days)	Standard Deviation Recovery Time (days)
Standard	110	217	34
New	100	186	29

N - O

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

We want to test the following hypotheses at the $\alpha = 0.05$ significance level.

$H_0: \mu_{diff} = 0$ days

$H_a: \mu_{diff} < 0$ days

where μ_{diff} = mean difference in recovery times between those who receive the new procedure and those who receive the standard procedure ($\mu_{diff} = \mu_{new} - \mu_{standard}$).

We will conduct a two sample t-test if conditions are met.

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

 4B2

If you need more room for your work in part (b), use the space below.

-The normal condition is met by the central Limit Theorem because $n = 110 > 30$ and $n = 100 > 30$.

-The random condition is met as given because patients were randomly assigned to treatments.

-The independent condition is met because $10(210) = 2100 <$ total number of patients needing surgery to repair a torn ACL.

Test Statistic:

$$t = \frac{(186 - 217) - 0}{\sqrt{\frac{(29)^2}{100} + \frac{(34)^2}{110}}} = -7.127$$

$t = -7.127$ $p = p(t \leq -7.127) = 8.36 \times 10^{-12}$

since the p-value of 8.36×10^{-12} is less than any reasonable α level, we reject H_0. The data provide convincing evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure.

GO ON TO THE NEXT PAGE.

4C1 4C1

4. The anterior cruciate ligament (ACL) is one of the ligaments that help stabilize the knee. Surgery is often recommended if the ACL is completely torn, and recovery time from the surgery can be lengthy. A medical center developed a new surgical procedure designed to reduce the average recovery time from the surgery. To test the effectiveness of the new procedure, a study was conducted in which 210 patients needing surgery to repair a torn ACL were randomly assigned to receive either the standard procedure or the new procedure.

(a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

No, correlation doesn't prove causation. They could conclude that the new procedure will most likely reduce recovery time compared to the standard procedure, for patients similar to those in the study, but they can't prove that it causes a reduction in recovery time.

(b) Summary statistics on the recovery times from the surgery are shown in the table.

Type of Procedure	Sample Size	Mean Recovery Time (days)	Standard Deviation Recovery Time (days)
Standard	110	217	34
less New	100	186	29

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

Assumptions: SRS, independence, $n_1p_1 \geq 10$, $n_1q_1 \geq 10$, $n_2p_2 \geq 10$, $n_2q_2 \geq 10$
sample size < 10% population

$H_0: \mu_{new} = \mu_{standard}$

$H_a: \mu_{new} < \mu_{standard}$

$$P\left(z < \frac{186 - 217}{\sqrt{\frac{29^2}{100} + \frac{34^2}{110}}}\right)$$

$P(z < -7.127)$

$P \approx 0$

GO ON TO THE NEXT PAGE.

If you need more room for your work in part (b), use the space below.

Since p is so small, we can reject H_0. This means that there is convincing statistical evidence to conclude that those who receive the new procedure will likely have less recovery time from the surgery, on average, than those who receive the standard procedure, for ~~the~~ patients similar to those in the study.

Question 4

Overview

The primary goals of this question were to assess a student's ability to (1) determine whether a cause-and-effect conclusion can be made based on how a study was conducted and (2) set up, perform, and interpret the results of a hypothesis test, in the context of the problem.

Sample: 4A
Score: 4

In part (a) component 1 of section 1 is satisfied because the response states that there is a causal relationship. Component 2 is satisfied because the causal relationship is justified based on the random assignment of the treatments. Component 3 is satisfied because the conclusion is given in context by using the word "patients." All three components of section 1 are satisfied. Section 1 was scored as essentially correct. In part (b) the parameters are correctly defined by using standard notation, μ, for a population mean. The generic subscripts 1 and 2 are used to denote the two different means; however, definitions are provided that clearly indicate that the subscript 1 refers to the standard procedure, and the subscript 2 refers to the new procedure. The definitions neglect to identify the parameters as the "population" mean, but this was overlooked in scoring because commonly accepted notation is used. Component 1 of section 2 is satisfied. Component 2 of section 2 is satisfied because the hypotheses indicate equality in the null and the correct direction in the alternative. Component 3 of section 2 is satisfied because the two-sample *t*-test is correctly identified by stating the correct formula for the test statistic, and the name of the test is given. Only the name or the formula is required to satisfy component 3. Component 4 of section 2 is satisfied because the correct test statistic, 7.127, is stated. All four components of section 2 are satisfied, and section 2 was scored as essentially correct. In part (b) the response makes reference to a correct *p*-value for a difference in means that is consistent with the test statistic and alternative hypothesis, satisfying component 1 of section 3. The response uses the small *p*-value to justify the conclusion that there is evidence to reject the null, satisfying component 2 of section 3. The response provides a correct conclusion by stating that there is evidence to support the alternative hypothesis, and this conclusion is stated in the context of the study. All three components of section 3 are satisfied, and section 3 was scored as essentially correct. Because the response was scored as essentially correct in three sections, the conditions of the test must be checked. Two conditions are required: the treatments must be randomly assigned, and the number of subjects in each treatment group must be large. The response clearly states that the treatments are randomly assigned. The response states that both samples are greater than 30; therefore, the "sampling distributions [of the difference in the sample means] are approximately normal." The response correctly notes that the condition that the sample size can be no larger than 10 percent of the population is unnecessary for this study because the study is a randomized experiment, not a random sample. Because three sections were scored as essentially correct, and conditions were stated and checked, the response earned a score of 4.

Sample: 4B
Score: 3

In part (a) component 1 of section 1 is satisfied because the response correctly states that it is reasonable to make a causal conclusion. Component 2 of section 1 is satisfied because the response justifies the causal relationship based on the random assignment of patients to procedures. Although not necessary, the response provides a very nice explanation of why the random assignment of patients to procedures reduces the possible effects of uncontrolled variables to allow for the conclusion of a causal relationship. Component 3 of section 1 is satisfied because the response is stated in context. All three components in section 1 are satisfied, and section 1 was scored as essentially correct. In part (b) component 1 of section 2 is not satisfied because the parameters are not correctly defined. The response defines the population mean for the difference in means for two dependent samples, which is not correct. Component 2 of section 2 is satisfied because the hypotheses indicate equality in the null and the correct

Question 4 (continued)

direction in the alternative. The direction is able to be determined because the response stated that $\mu_D = \mu_{new} - \mu_{standard}$ and $\mu_D < 0$, so it is clear that the alternative is testing that the mean for the new procedure is less than the mean for the standard procedure. Component 3 of section 2 is satisfied because the response states that a two-sample *t*-test will be conducted and uses the correct formula for the test statistic. Only the name or the formula is required to satisfy component 3. Component 4 of section 2 is satisfied because the response states the correct test statistic for a difference in means. All four components of section 2 are satisfied, and section 2 was scored as partially correct. In part (b) component 1 of section 3 is satisfied because the response provides a correct *p*-value that is consistent with the test statistic and the alternative hypothesis. The conclusion is justified based on the small size of the *p*-value, satisfying component 2 of section 3. Component 3 of section 3 is satisfied because the conclusion is stated in the context of the study. All three components are satisfied and section 3 was scored as essentially correct. Because two sections were scored as essentially correct, and one section was scored as partially correct, the response earned a score of 3.

Sample: 4C
Score: 2

In part (a) the response states that it is incorrect to conclude a causal relationship; therefore, component 1 of section 1 is not satisfied. To be scored as partially correct, component 1 must be satisfied with weak justification. Therefore section 1 was scored as incorrect. In part (b) component 1 of section 2 is satisfied because the parameters are correctly identified using standard notation, μ_{new} and $\mu_{standard}$. Component 2 of section 2 is satisfied because the hypotheses indicate equality in the null and the correct direction in the alternative. Component 3 of section 2 is not satisfied because the response identifies the test as a two-sample *z*-test instead of a two-sample *t*-test. Component 4 of section 2 is satisfied because the correct test statistic for a difference in means is provided. Three of the four components of section 2 are satisfied, and section 2 was scored as partially correct. In part (b) component 1 of section 3 is satisfied because the response makes reference to an approximately correct *p*-value that is consistent with the test statistic and alternative hypothesis by stating that $P \approx 0$. Component 2 of section 3 is satisfied because the response refers to a small *p*-value and uses that as justification for evidence to support the alternative hypothesis. Component 3 is satisfied because the conclusion is correctly stated in context. All three components of section 3 are satisfied, and section 3 was scored as essentially correct. Because one section was scored as essentially correct, one section was scored as partially correct, and one section was scored as incorrect, the response earned a score of 2.

2018

AP® CollegeBoard

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free Response Question 5

☑ Scoring Guideline

☑ Student Samples

☑ Scoring Commentary

Question 5

Intent of Question

The primary goals of this question were to assess a student's ability to (1) determine which of two histograms represents data with a larger median; (2) calculate the mean of a combined data set when the separate means and sample sizes are known; and (3) calculate the probability that an individual randomly chosen from a finite population will have a value within one standard deviation of the mean, when provided with values for the mean, standard deviation, and all members of the population.

Solution

Part (a):

The median teaching year for High School A is any value with 100 data values at or below it and 100 data values at or above it. The median teaching year for High School B is the 111th value in the ordered list of values. For High School A the median is in the interval that starts at 7 and ends just before 10, because there are only 94 data values below 7 and 106 data values of at least 7. Therefore the median cannot be less than 7. For High School B the median is in the interval that starts at 4 and ends just before 7 because there are more than half (113) of the data values less than 7. Therefore the median must be less than 7. So High School A must be the one with a median of 7, and High School B must be the one with a median of 6.

Another way to determine which school has the median of 7 is to notice that the distribution for High School B is highly skewed to the right, whereas the distribution for High School A is bimodal with a few possible outliers on the right. A distribution that is highly right-skewed is likely to have a substantially larger mean than median. The mean of both distributions is given as 8.2 years, so it makes sense that the highly right-skewed distribution (High School B) is the one with the bigger gap between the mean and median and, therefore, the one with the lower median of 6.

Part (b):

The mean for the original 200 teachers was given as 8.2 years, and the mean for the additional 18 teachers is 2.5 years. Therefore the mean for the combined data set is:

$$\frac{(200)(8.2) + (18)(2.5)}{200 + 18} = \frac{1{,}640 + 45}{218} \approx 7.73 \text{ years.}$$

Part (c):

The interval mean plus or minus 1 standard deviation on either side of the mean is 8.2 ± 7.2, or from 1.0 year to 15.4 years. Because teaching year is recorded as an integer, the interval includes teaching years 1 to 15. The number of teachers in that interval can be found by adding the heights of the five bars in the histogram for the intervals from 1 to 16, which includes $79 + 34 + 28 + 29 + 19 = 189$. Therefore the probability is $\frac{189}{221} \approx 0.8552$.

Question 5 (continued)

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. States that the median is 6 for High School B and the median is 7 for High School A.
2. Provides a reasonable explanation of how the decision was made.
3. Provides the definition of the median or explicitly applies the definition of a median as a criterion in reaching their decision.

OR

Essentially correct (E) if the response satisfies the following three components:

1. States that the median is 6 for High School B and the median is 7 for High School A.
2. States that High School B shows a skewed distribution (or High School A shows a less skewed distribution).
3. Provides a reasonable explanation of how the more skewed distribution (High School B) would be the one with a larger separation between the mean and median.

Partially correct (P) if the response satisfies the first component and only one of the other two components required for E.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: An incorrect statistical statement in the response will result in E being lowered to P, but not P being lowered to I. For example,

- If either distribution is described as left skewed, normal, or approximately normal;
- If the discussion would indicate a median different than 7 for High School A or a median different than 6 for High School B.

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following two components:

1. The correct answer that the mean is 7.73.
2. Enough work to show that the answer was obtained as a weighted average of the two individual means.

Partially correct (P) if the response satisfies only one of the two components.

Incorrect (I) if the response does not satisfy the requirements for E or P.

Question 5 (continued)

Part (c) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Calculates that the appropriate interval is 1 to 15.4 or 1 to 15 teaching years.
2. Correctly sums the counts of data values in the numerator based on the intervals provided.
3. Computes the probability using 221 as the denominator.

Partially correct (P) if the response satisfies only two of the three components;
OR
if the response reports the correct probability (0.8552) without supporting work.

Incorrect (I) if the response satisfies at most one of the three components.

Notes:

- If the response attempts to use the Empirical Rule or normal distribution to provide the desired probability, the response is scored I.
- If an incorrect count is shown in component 2, for instance by including the interval from 16 to 19, then component 3 is satisfied if that incorrect count is divided by 221 to find the reported probability.
- It is acceptable if the count is slightly off because of difficulty reading the exact heights of the bars in the histogram.
- If only one of component 2 or component 3 is missing, but the correct probability (0.8552) is reported, the response can be scored E.
- If the response recognizes that all values in the histogram bins up to 16 fall within one standard deviation of the mean and reports the interval as 1 to 16, component 1 is satisfied.

Question 5 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

OR

Part (a) essentially correct and two parts partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

Part (b) or part (c) essentially correct and one or two parts partially correct

OR

Three parts partially correct

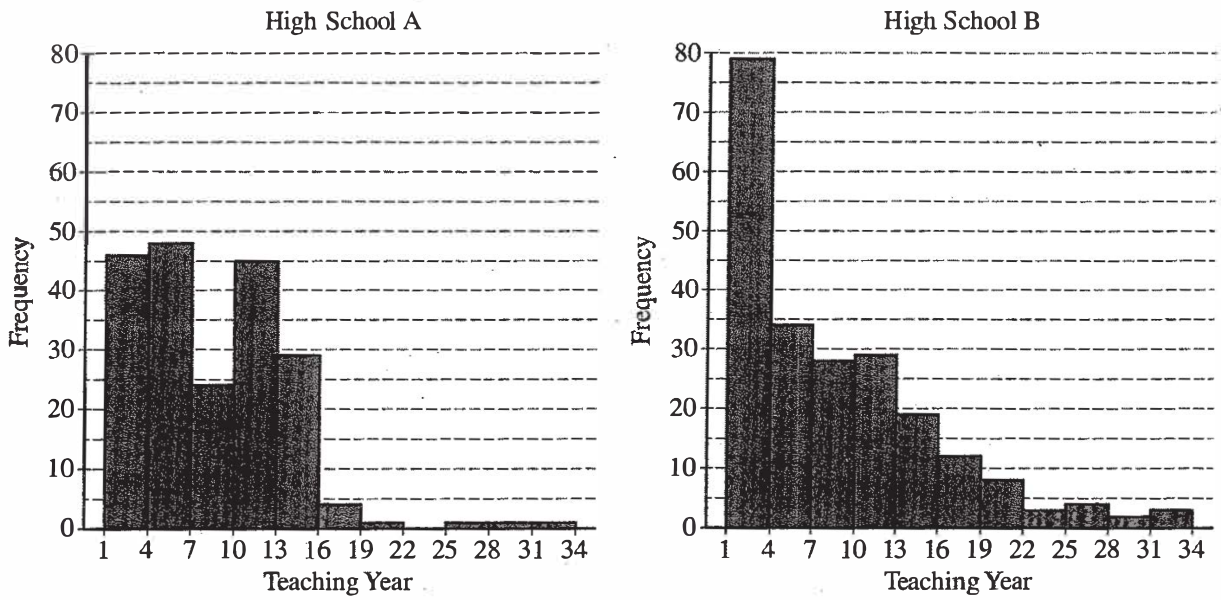
1 Minimal Response

One part essentially correct

OR

No parts essentially correct and one or two parts partially correct

5. The following histograms summarize the teaching year for the teachers at two high schools, A and B.



Teaching year is recorded as an integer, with first-year teachers recorded as 1, second-year teachers recorded as 2, and so on. Both sets of data have a mean teaching year of 8.2, with data recorded from 200 teachers at High School A and 221 teachers at High School B. On the histograms, each interval represents possible integer values from the left endpoint up to but not including the right endpoint.

(a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

High school A has 200 teachers so its median should be the mean of the 100th and 101th values when the values are ordered. Based on the chart, about <100 have values below 7 but >100 have values below 10 the median must be 7 at high school A.

Similarly in high school B the median of the 221 teachers will be the 111th value when ordered.

The chart shows there are <111 below 4 and >111 below 7. The median must be 6 at high school B.

GO ON TO THE NEXT PAGE.

(b) An additional 18 teachers were not included with the data recorded from the 200 teachers at High School A. The mean teaching year of the 18 teachers is 2.5. What is the mean teaching year for all 218 teachers at High School A?

The sum of the 200 teachers in the given chart is $\mu_1 n_1 = 8.2 \cdot 200 = 1640$.
The added teachers' sum would be $\mu_2 n_2 = 2.5 \cdot 18 = 45$
The total would be 1685 years for 218 teachers.
The mean would be

$$\mu_{1,2} = \frac{1640 + 45}{200 + 18} \approx \boxed{7.73 \text{ years}}$$

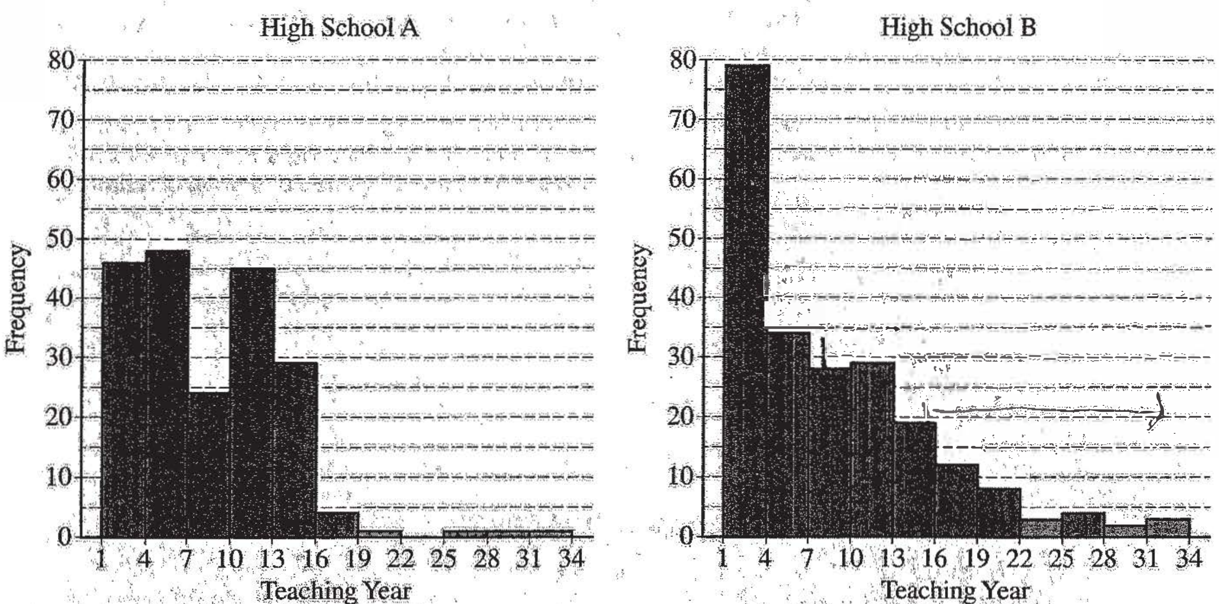
(c) The standard deviation of the teaching year for the 221 teachers at High School B is 7.2. If one teacher is selected at random from High School B, what is the probability that the teaching year for the selected teacher will be within 1 standard deviation of the mean of 8.2 ? Justify your answer.

Falling with 7.2 years of 8.2 and being integral means the value must be between $[1, 15]$, or equivalent $[1, 16)$.
This is the sum of values in the first 5 bars: $79 + 34 + 28 + 29 + 19 = 189$ values < 16
The proportion is $\frac{189}{221} \approx \boxed{0.855}$

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

5. The following histograms summarize the teaching year for the teachers at two high schools, A and B.



Teaching year is recorded as an integer, with first-year teachers recorded as 1, second-year teachers recorded as 2, and so on. Both sets of data have a mean teaching year of 8.2, with data recorded from 200 teachers at High School A and 221 teachers at High School B. On the histograms, each interval represents possible integer values from the left endpoint up to but not including the right endpoint.

(a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

By looking at the graphs, School B is skewed to the right much stronger than School A. When distributions are more skewed, they pull the mean towards the tail more. Since School B is more skewed, there's a greater difference between mean and median in School B. And since they both have the same mean, School B will have the lower median of 6 and School A will have the higher median of 7.

GO ON TO THE NEXT PAGE.

5B2

(b) An additional 18 teachers were not included with the data recorded from the 200 teachers at High School A. The mean teaching year of the 18 teachers is 2.5. What is the mean teaching year for all 218 teachers at High School A?

$$\frac{(200)(8.2) + (18)(2.5)}{218} = 7.73$$

with the 18 additional teachers, the mean teaching year decreases to 7.73.

(c) The standard deviation of the teaching year for the 221 teachers at High School B is 7.2. If one teacher is selected at random from High School B, what is the probability that the teaching year for the selected teacher will be within 1 standard deviation of the mean of 8.2 ? Justify your answer.

8.2 ± 7.2

$(1, 15.4)$ ← Interval the selected teacher ~~[illegible]~~ must fall in.

$1 - P(\text{year} > 15.4) = 1 - \frac{51}{221}$ ← There are about 51 teachers with teaching year above 15.4

$= 1 - .2308$

$\approx .7692$

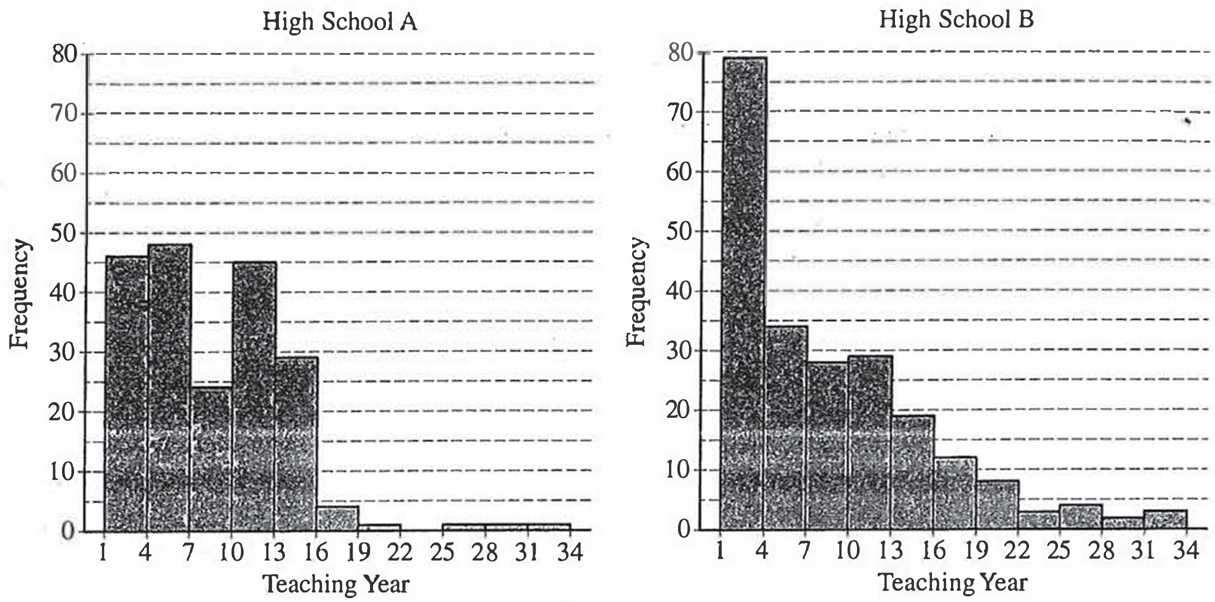
There is about a 76.92% chance that a random teacher will have a teaching year within 1 standard deviation of the mean of 8.2

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

5C1

5. The following histograms summarize the teaching year for the teachers at two high schools, A and B.



Teaching year is recorded as an integer, with first-year teachers recorded as 1, second-year teachers recorded as 2, and so on. Both sets of data have a mean teaching year of 8.2, with data recorded from 200 teachers at High School A and 221 teachers at High School B. On the histograms, each interval represents possible integer values from the left endpoint up to but not including the right endpoint.

(a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

High School A has the median teaching year of 7 and High School B has the median teaching year of 6. According to the graphs. High School A has 46+48=94 teachers with 1 to 6 teaching year. Since High School A has 200 teachers recorded the median is between the 100th and 101th teacher which both fall in the range of 7-9 teaching years. so High School A has median teaching year of 7. According to the graphs, High School B has 79+34=113 teachers with 1 to 6 teaching year. Since High School B has 221 teachers recorded. the median is the 111th teacher which falls in the range of 4-6 teaching years. so High School B has median teaching year of 6.

GO ON TO THE NEXT PAGE.

5C2

(b) An additional 18 teachers were not included with the data recorded from the 200 teachers at High School A. The mean teaching year of the 18 teachers is 2.5. What is the mean teaching year for all 218 teachers at High School A?

$$\mu_A = \frac{8.2 \times 200 + 18 \times 2.5}{218} = \frac{1685}{218} = 7.73 \text{ years.}$$

(c) The standard deviation of the teaching year for the 221 teachers at High School B is 7.2. If one teacher is selected at random from High School B, what is the probability that the teaching year for the selected teacher will be within 1 standard deviation of the mean of 8.2 ? Justify your answer.

$\mu = 8.2$ yrs . $\sigma = 7.2$ yrs

data within 1 standard deviation of the mean: (1 , 15.4)

normalcdf (1, 15.4, 8.2, 7.2) = 0.683

Since the graph is strongly skewed to the right, the predicted probability may be inaccurate

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

Question 5

Overview

The primary goals of this question were to assess a student's ability to (1) determine which of two histograms represents data with a larger median; (2) calculate the mean of a combined data set when the separate means and sample sizes are known; and (3) calculate the probability that an individual randomly chosen from a finite population will have a value within one standard deviation of the mean, when provided with values for the mean, standard deviation, and all members of the population.

Sample: 5A
Score: 4

In part (a) the response correctly identifies the location of the median for High School A as "the mean of the 100^{th} and 101^{th} values when the values are ordered" and the location of the median for High School B as "the 111^{th} value when ordered," which satisfies component 3. For High School A, the response indicates "about < 100 have values below 7 and > 100 have values below 10," which satisfies component 2. The response correctly concludes that "the median must be 7 at high school A," which satisfies component 1. For High School B, the response correctly counts that there are "< 111 [values] below 4 and >111 [values] below 7," which again satisfies component 2. The response correctly concludes that "The median must be 6 at high school B," which again satisfies component 1. Because the response satisfies all three components, part (a) was scored as essentially correct. In part (b) the response uses the correct weights to compute the total number of teaching years for the initial 200 teachers (1,640) and for the 18 additional teachers (45), which satisfies component 2. The response adds these values, divides the total number of teaching years by 218, and reports the correct mean, which satisfies component 1. Because the response includes both components, part (b) was scored as essentially correct. In part (c) the response provides two appropriate intervals. The response points out that because the data values are recorded as integers, if a value is within one standard deviation of the mean, the value will be in the interval [1, 15] or equivalently [1, 16), which satisfies component 1. The response states that the appropriate number of data values "is the sum of values in the first 5 bars" and computes the correct sum, which satisfies component 2. The sum is divided by the correct denominator (221), and the correct probability is reported, which satisfies component 3. Because the response includes all three components, part (c) was scored as essentially correct. Because three parts were scored as essentially correct, the response earned a score of 4.

Sample: 5B
Score: 3

In part (a) the response correctly identifies the median for the distribution of teaching years for High School A as 7 and the median for the distribution of teaching years for High School B as 6, which satisfies component 1. The response bases the decision on the distribution of High School B being "skewed to the right much stronger than school A," which satisfies component 2. The response satisfies component 3 when it points out that in a skewed distribution the mean will be pulled towards the tail and states that "Since school B is more skewed there is a greater difference between the mean and median in school B." Because the response includes all three components, part (a) was scored as essentially correct. In part (b) the response provides an equation with the correct weighted sum of means in the numerator divided by the total number of teachers and reports the correct mean, which satisfies component 1 and component 2. Because the response includes both components, part (b) was scored as essentially correct. In part (c) the response reports the correct interval supported by appropriate calculations, which satisfies component 1. The probability that the randomly selected teaching year will be more than one standard deviation from the mean is calculated and subtracted from one. In this calculation the denominator is correct, which satisfies component 3. However, the numerator is incorrect (51 instead of 32) because when summing frequencies for observations with "year >15.4" the response includes those values in the interval from 13 to 16; therefore,

Question 5 (continued)

component 2 is not satisfied. Because the response satisfies only two of the three components, part (c) was scored as partially correct. Because two parts were scored as essentially correct, and one part was scored as partially correct, the response earned a score of 3.

Sample: 5C
Score: 2

In part (a) the response correctly identifies the median for the distribution of teaching years for High School A as 7 and the median for the distribution of teaching years for High School B as 6, which satisfies component 1. For High School A, the response correctly counts that there are "94 teachers with 1 to 6 teaching year" and that the "median is between the 100th and 101th teacher" and concludes that the median is 7, which satisfies component 2 and component 3. For High School B, the response correctly counts that there are "113 teachers with 1 to 6 teaching year" and that "the median is the 111th teacher" and concludes that the median is 6, which again satisfies component 2 and component 3. Because the response includes all three components, part (a) was scored as essentially correct. In part (b) the response provides an equation with the correct weighted sum of means in the numerator divided by the total number of teachers and reports the correct mean, which satisfies component 1 and component 2. Because the response includes both components, part (b) was scored as essentially correct. In part (c) the response computes the correct interval, which satisfies component 1. The response uses the normal distribution and computes an incorrect probability of 0.683. The response recognizes that the solution may not be correct and comments that "Since the graph is strongly skewed to the right the predicted probability may be inaccurate." However, because the response uses the normal distribution, part (c) was scored as incorrect. Because two parts were scored as essentially correct, and one part was scored as incorrect, the response earned a score of 2.