STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II grade—25

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

6. The statistics department at a large university is trying to determine if it is possible to predict whether an applicant will successfully complete the Ph.D. program or will leave before completing the program. The department is considering whether GPA (grade point average) in undergraduate statistics and mathematics courses (a measure of performance) and mean number of credit hours per semester (a measure of workload) would be helpful measures. To gather data, a random sample of 20 entering students from the past 5 years is taken. The data are given below.

Successfully Completed Ph.D. Program

Student	A	B	C	D	E	F	G	H	I	J	K	L	M
GPA	3.8	3.5	4.0	3.9	2.9	3.5	3.5	4.0	3.9	3.0	3.4	3.7	3.6
Credit hours	12.7	13.1	12.5	13.0	15.0	14.7	14.5	12.0	13.1	15.3	14.6	12.5	14.0

Did Not Complete Ph.D. Program

Student	N	O	P	Q	R	S	T
GPA	3.6	2.9	3.1	3.5	3.9	3.6	3.3
Credit hours	11.1	14.5	14.0	10.9	11.5	12.1	12.0

The regression output at the top of the next page resulted from fitting a line to the data in each group. The residual plots (not shown) indicated no unusual patterns, and the assumptions necessary for inference were judged to be reasonable.

GO ON TO THE NEXT PAGE.

Successfully Completed Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	23.514	1.684	13.95	0.000
GPA	–2.7555	0.4668	–5.90	0.000
S = 0.5658		R-Sq = 76.0%		

Did Not Complete Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	24.200	3.474	6.97	0.001
GPA	–3.485	1.013	–3.44	0.018
S = 0.8408		R-Sq = 70.3%		

(a) Use an appropriate graphical display to compare the GPA's for the two groups. Write a few sentences commenting on your display.

(b) For the students who successfully completed the Ph.D. program, is there a significant relationship between GPA and mean number of credit hours per semester?

Give a statistical justification to support your response.

(c) If a new applicant has a GPA of 3.5 and a mean number of credit hours per semester of 14.0, do you think this applicant will successfully complete the Ph.D. program? Give a statistical justification to support your response.

END OF EXAMINATION

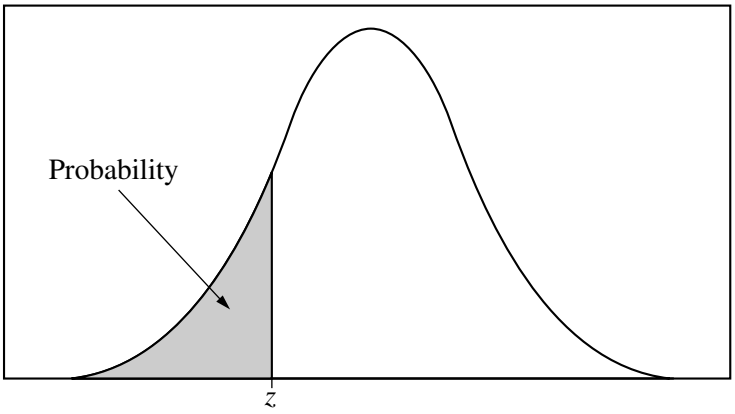

Table entry for z is the probability lying below z.

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

Question 6 (cont'd.)

Note: (a) A graph with no comment is incorrect.
(b) A correct comment that clearly indicates what graphical display was used, but the display is not shown, is partially correct.

Part (b):

Part 1: States a correct pair of hypotheses in symbols or in words

$$H_0: \beta = 0 \qquad \text{OR} \qquad H_0: \rho = 0$$
$$H_a: \beta \neq 0 \qquad\qquad H_a: \rho \neq 0$$

OR

H_0: There is no relationship between GPA and mean number of credit hours
H_a: There is a relationship between GPA and mean number of credit hours

Part 2: Identifies correct test by name or by formula and states that the appropriate assumptions were met.

t test for slope or linear regression t-test (LinReg t test) or gives t from the computer output.

$$t = \frac{\hat{\beta} - 0}{s_{\hat{\beta}}}$$

(Calculator output: t = -5.903, p = 0.0001026)

Problem states that assumptions necessary for inference were judged to be reasonable. This should be mentioned, but then there is no need to go further with checking of assumptions. (Correct assumptions are: random sample, the residuals from the line are normally distributed with mean zero and constant variance.)

Part 3: Correct mechanics, including value of test statistic and P-value (or rejection region)

The computer output provided: t = -5.90 P-value = .000

Part 4: States a correct conclusion in the context of the problem, linked to the P-value or rejection region.

Since the P-value is 0.000, reject the null hypothesis. For students who successfully completed the program, there is a significant relationship between GPA and mean number of credit hours per semester.

Note: If conclusion immediately follows P-value, the linkage is implied.

Note: If the student just appeals to the P-value and then goes directly to a correct conclusion, part (b) should be scored as partially correct.

Note: If the student appeals to high correlation as evidence of a significant relationship, part (b) should be scored as incorrect.

Question 6 (cont'd.)

Part (b) is

Essentially correct if

All 4 parts of the appropriate test are correct.

Part 1:	Hypotheses
Part 2:	Identifies test
Part 3:	Mechanics
Part 4:	Conclusion

Partially correct if

2 or 3 parts of the test are correct

OR

Student does not carry out a formal test, but correctly reasons from P-value on output to a correct conclusion.

OR

Student indicates that the assumptions necessary for inference have been met or that a linear equation is a good fit, and appeals to the high correlation.

Part (c)

Part (c) must be scored holistically.

To be considered essentially correct, it must be treated as a bivariate problem—both variables (GPA and credit hours) must be used. Also, the possibility of inclusion must be checked for both groups. The correct prediction (successful) must follow from the student's analysis.

The solution should be considered partially correct if it is treated as a bivariate problem but the communication is weak or the statistical justification is weak or difficult to follow.

The solution should be considered partially correct if it is treated as a bivariate problem but only one group (successful or unsuccessful) is used and the point is never considered in relation to the other group.

If a decision is made based on looking at either of the two variables individually, part (c) is incorrect.