FREE RESPONSE

Questions 1-3

Spend about 45 minutes on this part of the exam.

2003 AP® STATISTICS FREE-RESPONSE QUESTIONS

5. A random sample of 200 students was selected from a large college in the United States. Each selected student was asked to give his or her opinion about the following statement.

> "The most important quality of a person who aspires to be the President of the United States is a knowledge of foreign affairs."

Each response was recorded in one of five categories. The gender of each selected student was noted. The data are summarized in the table below.

	Response Category				
	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
Male	10	15	15	25	25
Female	20	25	25	25	15

Is there sufficient evidence to indicate that the response is dependent on gender? Provide statistical evidence to support your conclusion.

H_0: Response and gender are independent

H_A: Response and gender are not independent.

$\alpha = .05$

I will perform a χ^2 test for independence

The sample was randomly selected and all expected counts were at least 5. (The smallest was 18.) So the conditions were satisfied.

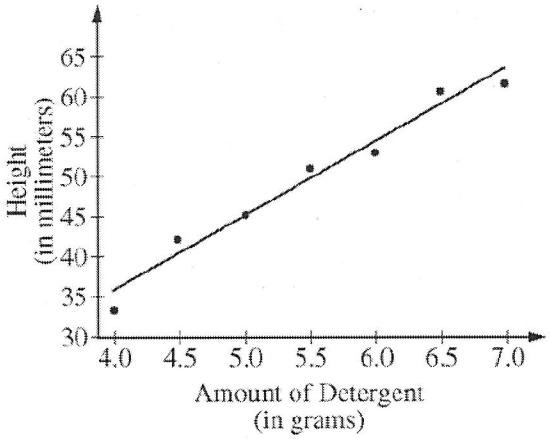
$\chi^2 = 8.9226$ and $p = .0631$

With a p-value of $.0631 > \alpha = .05$, we fail to reject H_0. There is not enough evidence to conclude that response and gender are not independent.

2006 AP® STATISTICS FREE-RESPONSE QUESTIONS

2. A manufacturer of dish detergent believes the height of soapsuds in the dishpan depends on the amount of detergent used. A study of the suds' heights for a new dish detergent was conducted. Seven pans of water were prepared. All pans were of the same size and type and contained the same amount of water. The temperature of the water was the same for each pan. An amount of dish detergent was assigned at random to each pan, and that amount of detergent was added to the pan. Then the water in the dishpan was agitated for a set amount of time, and the height of the resulting suds was measured.

A plot of the data and the computer output from fitting a least squares regression line to the data are shown below.



Predictor	Coef	SE Coef	T	P
Constant	−2.679	4.222	−0.63	0.554
Amount	9.5000	0.7553	12.58	0.000

S = 1.99821 R-Sq = 96.9% R-Sq(adj) = 96.3%

(a) Write the equation of the fitted regression line. Define any variables used in this equation.

$\widehat{\text{Height}} = -2.679 + 9.500\,(\text{Amount of Detergent})$

(b) Note that $s = 1.99821$ in the computer output. Interpret this value in the context of this study.

$s = 1.99821$ is the standard deviation of the residuals. It measures how much variability there is in the height of soapsuds vertically to the regression line.

(c) Identify and interpret the standard error of the slope.

$se(b_1) = .7553$ mm per gram. This estimates the variability in the sampling distribution of the estimated slope.

(This one is challenging)

2007 Form A Question 6

A study was designed to explore subjects' ability to judge the distance between two objects placed in a dimly lit room. The researcher suspected that the subjects would generally overestimate the distance between the objects in the room and that this overestimation would increase the farther apart the objects were.

The two objects were placed at random locations in the room before a subject estimated the distance (in feet) between those two objects. After each subject estimated the distance, the locations of the objects were rerandomized before the next subject viewed the room.

After data were collected for 40 subjects, two linear models were fit in an attempt to describe the relationship between the subjects' perceived distances (y) and the actual distance, in feet, between the two objects.

Model 1: $\hat{y} = 0.238 + 1.080 \times$ (actual distance)

The standard errors of the estimated coefficients for Model 1 are 0.260 and 0.118, respectively.

Model 2: $\hat{y} = 1.102 \times$ (actual distance)

The standard error of the estimated coefficient for Model 2 is 0.393.

(a) Provide an interpretation in context for the estimated slope in Model 1.

1.080 estimates the average increase (in feet) in the perceived distance for each additional foot in actual distance between the objects.

(b) Explain why the researcher might prefer Model 2 to Model 1 in this context.

It makes more sense with a y-intercept of zero, meaning if two objects were placed side by side they should also have a perceived distance of zero.

(c) Using Model 2, test the researcher's hypothesis that in dim light participants overestimate the distance, with the overestimate increasing as the actual distance increases. (Assume appropriate conditions for inference are met.)

$H_0: \beta_1 = 1$

$H_A: \beta_1 > 1$

conditions are met, so I will perform a t-test for the slope.

$$t = \frac{b_1 - \beta_1}{s_b} = \frac{1.102 - 1}{.392} = 0.260$$

$\left(\frac{x-\mu}{\sigma}\right)$ ↗ with $40-1 = 39$ degrees of freedom

tcdf(.260, 9999, 39) = .398 = p-value

with a p-value $> \alpha = .05$, I fail to reject, there is not enough evidence that the participants overestimates increased. conclude

The researchers also wanted to explore whether the performance on this task differed between subjects who wear contact lenses and subjects who do not wear contact lenses. A new variable was created to indicate whether or not a subject wears contact lenses. The data for this variable were coded numerically (1 = contact wearer, 0 = noncontact wearer), and this new variable, named "contact," was included in the following model.

Model 3: $\hat{y} = 1.05 \times (\text{actual distance}) + 0.12 \times (\text{contact}) \times (\text{actual distance})$

The standard errors of the estimated coefficients for Model 3 are 0.357 and 0.032, respectively.

(d) Using Model 3, sketch the estimated regression model for contact wearers and the estimated regression model for noncontact wearers on the grid below.

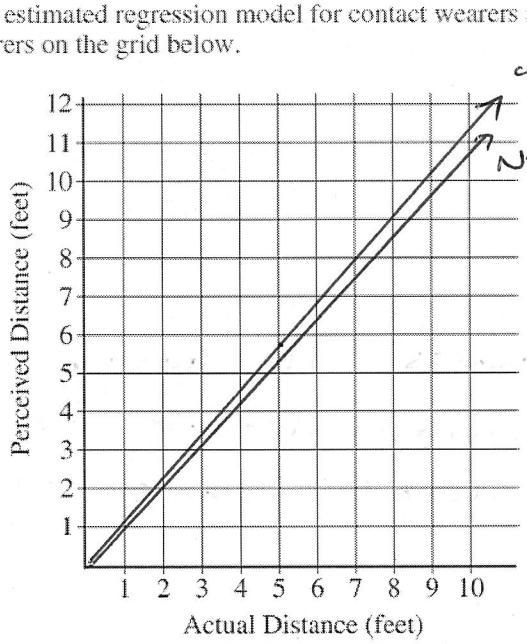
contacts:

$\hat{y} = 1.05x + .12x$

$= 1.17x$

No contacts:

$\hat{y} = 1.05(x)$



(e) In the context of this study, provide an interpretation of the estimated coefficients for Model 3.

Model 3 allows predictions for contact and non-contact wearers seperately. (This variable is taken into account). The 1.05 is the average perceived increase per foot of actual increase for non-contact wearers. The 0.12 estimates the additional increase for contact wearers.

AP® STATISTICS
2003 SCORING GUIDELINES

Question 5

Solution

Part 1: States a correct pair of hypotheses

H_o: Response and gender are independent
H_a: Response and gender are not independent
OR
H_o: There is no association between response and gender
H_a: There is an association between response and gender

Part 2: Identifies a correct test (by name or by formula) and checks appropriate conditions.

Chi-Square test (for independence)

$$\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp}$$

Conditions: Random sample and large sample size

Expected counts are

	Strongly Disagree	Somewhat Disagree	Neither Agree or Disagree	Somewhat Agree	Strongly Agree
Male	13.5	18.0	18.0	22.5	18.0
Female	16.5	22.0	22.0	27.5	22.0

All expected counts are greater than 5 (or 10), so the sample size is large enough for the Chi-Square test to be appropriate.

(Or, all expected counts are ≥ 1, and no more than 20% of expected counts < 5.)

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

$\chi^2 = 0.907+0.500+0.500+0.278+2.722+0.742+0.409+0.409+0.227+2.227 = 8.921$

df = 4 P-value = 0.063

(Or, using tables, $0.05 <$ P-value < 0.10, or rejection regions: $\alpha = 0.05$ is 9.48, $\alpha = 0.01$ is13.27)

Question 5 (cont'd)

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

Because P-value > selected α (or because χ^2 is not in the rejection region, or because the P-value is large), fail to reject the null hypothesis. There is not sufficient evidence to conclude that response is dependent on gender (or that response and gender are not independent, or that response and gender are associated)

OR

Because results this extreme would occur about 6 times in 100 by chance alone, there is marginal evidence to reject the null hypothesis and conclude that there is an association between response and gender.

Scoring

Note that the solution has 4 parts and each part is either correct or incorrect. No partial credit is given for individual parts.

4 **Complete Response**
Four parts are correct.

3 **Substantial Response**
Three parts are correct.

2 **Developing Response**
Two parts are correct.

1 **Minimal Response**
Only one part is correct.

AP® STATISTICS
2006 SCORING GUIDELINES

Question 2

Intent of Question

The primary goal of this question is to assess a student's ability to identify the estimated regression line and to identify and interpret important statistics from regression output provided by statistical software in the context of a practical problem.

Solution

Part (a):

The regression line is $\hat{y} = -2.679 + 9.5x$, where $\hat{y}$ represents the estimated (or predicted) mean height of the soapsuds and x represents the amount of detergent added to the pan.

Part (b):

The value $s = 1.99821$ mm is the standard deviation of the residuals. This statistic measures a typical amount of variability in the vertical distances from the observed height of the soapsuds to the regression line.

OR

The value $s = 1.99821$ mm is a measure of variation in the height of soapsuds for a given amount of detergent.

Part (c):

The standard error of the estimated slope parameter is 0.7553 mm per gram. Thus, the standard deviation of the estimated slope for predicting the height of soapsuds by using an amount of detergent is estimated to be 0.7553 mm per gram. This value estimates the variability in the sampling distribution of the estimated slope (i.e., how much we would expect sample slopes to vary from experiment to experiment).

Scoring

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

Part (a) is essentially correct (E) if the least squares regression line is correctly identified and the variables are correctly defined.

Part (a) is partially correct (P) if:

the least squares regression line is correctly identified and either of the two variables are not correctly defined;

OR

the least squares regression line is not presented using estimated or predicted notation, or $\hat{y}$, AND both variables are correctly defined;

OR

only one of the two values is correctly identified from the table and both variables are correctly defined.

Part (a) is incorrect (I) if the least squares regression line is incorrectly identified or not identified, and the variables are not correctly defined.

Question 2 (continued)

Notes:

- If y is identified as the height of the soapsuds and x is identified as the amount of detergent, then the student should get credit for defining the variables. However, y must be identified as an estimated height somewhere in the student response in order to get this part essentially correct.
- If $\hat{y}$ (or *estimated height*) is used to specify the regression line and y is identified as the height, the response should be scored as essentially correct. For example, a response of estimated height $= -2.679 + 9.5$amount, where y represents the height of suds an x represents the amount of detergent should be scored essentially correct.
- If the regression line is specified using y (or height) instead of $\hat{y}$ (or estimated height), the response is scored as partially correct. For example, a response of $y = -2.679 + 9.5x$ where y = height of suds and x = amount of detergent should be scored as partially correct.
- If the estimates of the intercept and slope are reversed ($\hat{y} = 9.5 - 2.679x$), then the response should be scored as incorrect.

Part (b) is essentially correct (E) if the standard deviation is correctly interpreted in the context of this study.

Part (b) is partially correct (P) if the standard deviation is correctly interpreted in general terms without context.

Part (b) is incorrect (I) if the response indicates that s is any other standard deviation (e.g., univariate distribution of y), or slope.

Notes:

- If s is interpreted as the estimated standard deviation of the differences between the observed values for the height of soapsuds and the values predicted from the regression line, the response should be scored essentially correct.
- If s is interpreted as a "typical" prediction error for estimating height from the amount of detergent, then the response should be scored essentially correct.

Part (c) is essentially correct (E) if the standard error is identified and interpreted correctly.

Part (c) is partially correct (P) if standard error is identified but interpretation is weak (e.g., the standard error is a standard deviation of the slope). The major idea of sampling variability is not included.

Part (c) is incorrect (I) if the standard error is not correctly identified, identified with no interpretation, or an incorrect interpretation is provided.

Question 2 (continued)

4 Complete Response

All 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 two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct and either zero or one part partially correct

OR

No parts essentially correct and two parts partially correct

AP® STATISTICS
2007 SCORING GUIDELINES

Question 6

Intent of Question

This question was designed to evaluate a student's ability to make inferences for simple linear regression models. Interpreting model parameters and comparing and contrasting different models are important skills that are also being assessed. Finally, a multiple regression model with a special variable, an indicator variable, is introduced to investigate whether the relationship between the predictor and response variable differs for two different groups of people. Students are asked to sketch the estimated line for both groups and interpret the estimated parameters in the multiple regression model.

Solution

Part (a):

The value 1.080 estimates the *average* increase (in feet) in the perceived distance for each additional foot in actual distance between the two objects.

Part (b):

The model with zero intercept makes more intuitive sense in this particular situation. If the two objects are placed side by side (so the actual distance is zero), then we would expect the subjects to say that the distance between the objects is zero.

Part (c):

Let β denote the true slope between the perceived distances and the actual distances. The researcher's hypothesis is equivalent to $\beta > 1$. Thus, we want to conduct a hypothesis test for the slope parameter.

Step 1: States a correct pair of hypotheses:

$H_0 : \beta = 1$
$H_a : \beta > 1$

Step 2: Correct mechanics, including the value of the test statistic and *p*-value (or rejection region).

This is a *t*-test of a slope.

$$t = \frac{b - \beta}{s_b} = \frac{1.102 - 1}{0.393} = 0.260$$

$$\text{df} = 40 - 1 = 39$$

$$p\text{-value} = P(t > .260) = 0.398$$

AP® STATISTICS
2007 SCORING GUIDELINES

Question 6 (continued)

Step 3: States a correct conclusion in the context of the problem, using the result of the statistical test.

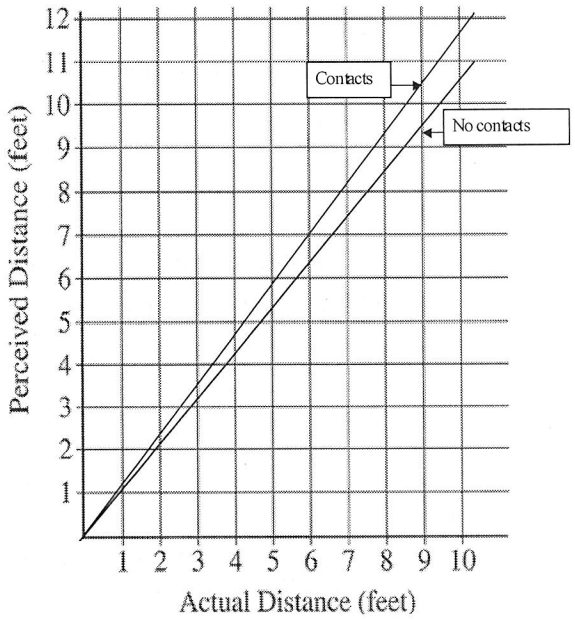
Since the p-value 0.398 is greater than 0.05, we cannot reject H_0. That is, we do not have statistically significant evidence to conclude that the subjects overestimate the distance with the magnitude of the overestimation increasing as the actual distance increases.

Part (d):

According to Model 3, the estimated models for the two groups are:

Contact wearers (*contact* = 1):
perceived distance = 1.05 (*actual distance*) + 0.12 (*actual distance*)
= 1.17 (*actual distance*)

Noncontact wearers (*contact* = 0):
perceived distance = 1.05 (*actual distance*)



Part (e):

Model 3 allows prediction of perceived distance separately for contact wearers and for noncontact wearers. The value of 1.05 estimates the average increase (in feet) in the perceived distance for each one-foot increase in actual distance for the population of noncontact wearers. The value of 0.12 estimates the *additional* increase (in feet) in the average perceived distance for each one-foot increase in actual distance for the contact wearers.

AP® STATISTICS
2007 SCORING GUIDELINES

Question 6 (continued)

Scoring

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

Parts (a) and (b) combined is scored as essentially correct (E) if both parts are correct.

Parts (a) and (b) combined is scored as partially correct (P) if:
one part is correct and the other part is incorrect;
OR
one part is correct and the other part is partially correct;
OR
both parts are partially correct.

Part (a) and (b) combined is scored as incorrect (I) if one part is partially correct.

Notes:

Part (a) is scored as partially correct if there is no word that makes it clear that 1.080 is not a deterministic increase.

Part (a) is scored as incorrect if the response:
- ignores the intercept and implies proportionality: for each foot of actual distance between the two objects, the subject perceives about 1.080 feet;
- consists of the equation rewritten in words.

Part (b)

Additional correct statement:
- The intercept is clearly not statistically significant, so the simpler model that includes only the slope is reasonable.

Partially correct statements:
- The SE for Model 2 is so large that Model 2 does not seem reasonable.
- The interpretation of the slope is straightforward if there is a 0 intercept: the percentage error is *slope* – 1 or 10.2 percent.
- The slope for Model 2 is farther above 1 than the slope for Model 1 and so more in line with the researcher's hypothesis.

Incorrect statements:
- Having one SE is better than having two.
- It is simpler/easier/shorter/more accurate to have just one coefficient.

AP® STATISTICS
2007 SCORING GUIDELINES

Question 6 (continued)

Part (c) is scored as:

Essentially correct (E) if three steps are correct.

Partially correct (P) if two steps are correct.

Incorrect (I) if one step is correct.

Notes:

- Hypotheses: the hypotheses step is incorrect if the alternative hypothesis is two-sided, or if the null hypothesis is $\beta = 0$. (It is not necessary to define β.)
- Computation: if the computation includes division by $\sqrt{40}$, the computation step is incorrect.
- Conclusion: a conclusion with no context is incorrect.

Part (d) is scored as essentially correct (E) if both estimated regression lines are graphed correctly and at least one is labeled.

Part (d) is scored as partially correct (P) if:

- the lines are graphed correctly but neither is labeled;
  OR
- the graphs consist of unconnected dots.

Part (d) is scored as incorrect (I) if:

- the two lines on the grid have the same slope;
  OR
- one line is plotted correctly and one line is not.

Part (e) is scored as essentially correct (E) if the response includes a correct interpretation of the estimated coefficients, 1.05 and 0.12. Unlike in part (a) there is no *y*-intercept, so this statement is correct: "For each foot of actual distance between the two objects, a noncontact wearer perceives about 1.05 feet, and a contact wearer will perceive about an additional 0.12 feet."

Part (e) is scored as partially correct (P) if:

- the response includes a correct interpretation of just one of the two coefficients;
  OR
- the response includes a correct interpretation of 1.05 and 1.05 + 0.12 = 1.17 but doesn't include a separate interpretation of 0.12;
  OR
- no numbers are mentioned, but it is made clear that both groups overestimate the distance *AND* that contact wearers overestimate more than do noncontact wearers.

Part (e) is scored as incorrect (I) if:

- the response says only that 1.05 and 0.12 are "slopes of regression lines";
  OR
- only the SEs of the coefficients, 0.357 and 0.032, are interpreted.

AP® STATISTICS
2007 SCORING GUIDELINES
Question 6 (continued)

Each essentially correct (E) response counts as 1 point; each partially correct (P) response counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.