

Psych 5741: Final Exam
Fall 1995

Question A

Janet Helms (1990) proposed a 5 stage model of African American racial identity development. This model was designed to outline the process by which an African American individual develops, if he/she ever does develop, a transcendent racial identity. Another researcher is interested in examining a variety of different factors which may affect this process in hopes of beginning to determine what factors, individually or in concert, are involved in racial identity development. To accomplish this purpose, the researcher has collected data on the following variables for 300 African American individuals.

IDENTITY: an individual's stage of racial identity development (on a 1-5 scale, where 1 represents Pre-encounter, the stage of an individual who denies their racial identity or who is ignorant of racial issues and has never thought about such issues and their impact on every day life, and 5 represents Internalization/Commitment, the stage of an individual who has dealt with and come to terms with racial issues and transcended them in forming his/her ethnic identity. Progression is seen as moving from stage 1 to stage 5.)

GENDER: an individual's gender (0-Male, 1-Female)

AGE: an individual's age (in years)

FAMILY: the level of support an individual's family offers him/her (on a 1-5 scale, where 1 is little family support and 5 is a high level of family support)

AREA: the percentage of African Americans living in the community where the individual resides

FRIENDS: the number of African American friends the individual has

ACTIVISM: the level of political activism an individual engages in (on a 1-5 scale, where 1 is no activism and 5 is a very high level of activism)

The researcher is interested in a variety of different questions and would like your help in developing the model comparisons needed to answer each of his questions. For each of the following questions, please provide Models A & C, the null hypothesis, PA-PC, and n-PA. Unless otherwise instructed, use the simplest model comparison possible.

1) Is the average African American in this study in stage 3 or higher?

2) Does an individual's family support significantly impact his/her level of racial/ethnic identity over and above the effects of gender and age?

3) It has been suggested that the more friends of one's race one has, the more racially identified one will be. Is this true of the current data?

4) Does having more African American friends impact a male's ethnic/racial identity more than a female's?

5) Research has suggested that political activism is a key component of an African American's racial identity. However, some researchers have qualified this by saying that those who are too radical (in this case, those who engage in too much activism) will be in a lower stage than individuals who engage in only a moderate amount of activism. Are the latter researchers' contentions accurate given the current data?

6) A researcher believes, in the context of the previous model, that the optimum level of activism is 4 in terms of maximizing racial identity stage. But that higher levels of activism become counter productive. Is 4 the optimum level of activism?

7) In the context of the previous model, does the mean of racial identity equal 4.5 for those with an activism of 4?

Question B

This problem is based on a business example as a thank you to the business students in the class who have tolerated the many psychological examples throughout the semester.

A large national firm with many retail outlets wants to investigate the relationship between each store's advertising budget and profit. In addition, the data analyst has information about the size of the store and the type (either electronics or major appliances). For each graph below, specify the **simplest** model that could be used to describe the relationship as depicted.

Express your models using 's (i.e., don't try to guestimate any parameter values) and use the following abbreviations for the variables:

P = Profits (in 1000s of dollars)

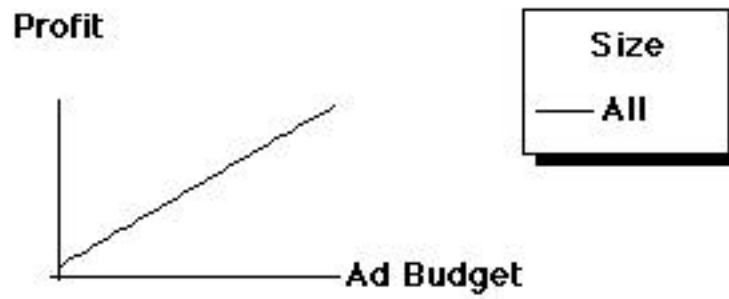
A = Advertising Budget (in 1000s of dollars)

S = Size (1000s of square feet)

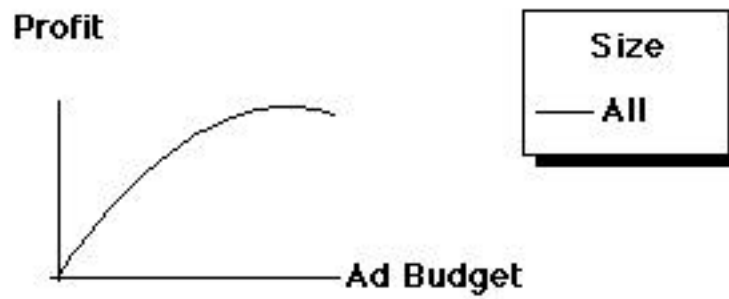
T = Type (electronics or appliances)

Note that Size is a continuous variable, but representative lines (or curves) are drawn separately for "small" (S) and "large" (L) retail outlets. (Note: we are not asking for model comparisons; rather, we are asking for the one model that corresponds to each graph.)

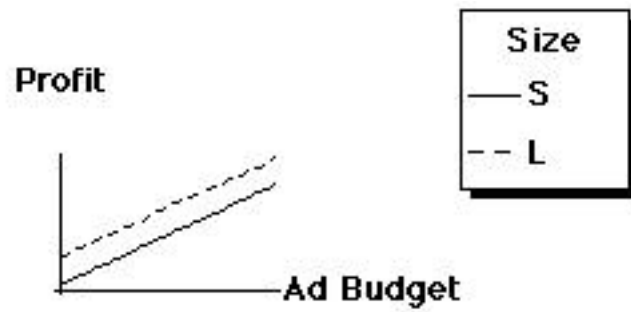
1.



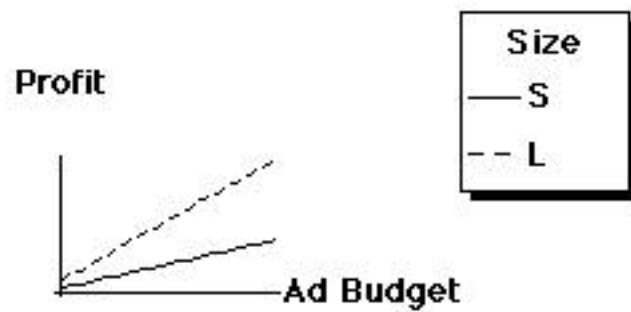
2.



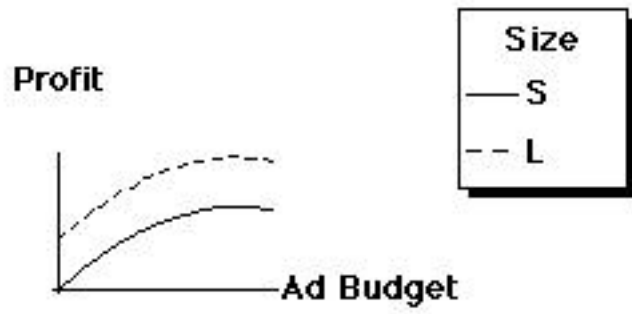
3.



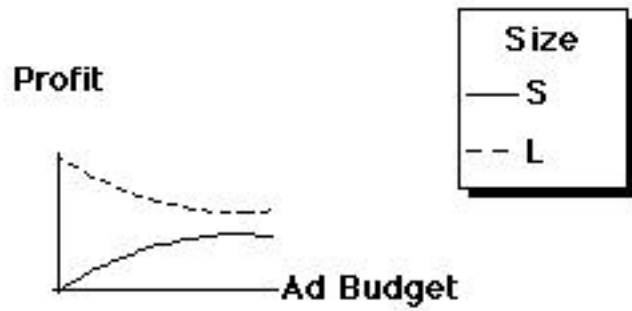
4.



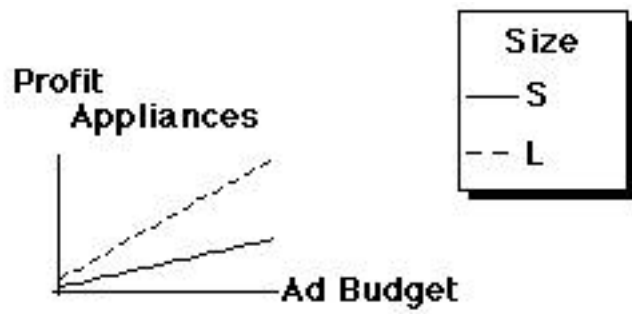
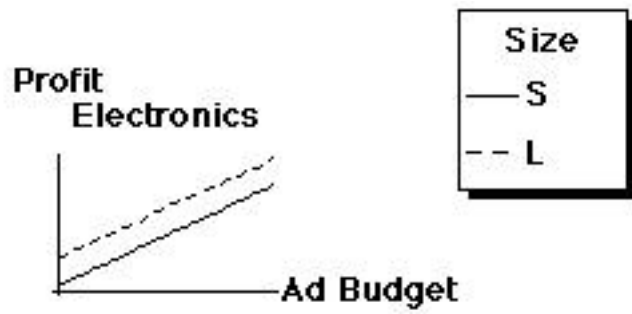
5.



6.



7.



8. Although we finished the term with very sophisticated models involving interactions and polynomials, we shouldn't forget that simple linear regression is still our "bread-and-butter" procedure. Suppose that the data analyst found that the model $P = 20 + 2.5 A$ had $SSE = 5,236$ for a group of 25 stores and that the variance of P for those same 25 stores was 351. Report PRE and F^* for the appropriate test, make a statistical conclusion, provide interpretations of both parameters, and write a brief substantive conclusion. [Note: this model is fictitious; if only it were so easy to make money!]

9. In a more complete analysis, the data analyst obtains the printout on the next page from StatView (a statistical program for the Mac). Interpret each coefficient precisely and indicate whether each coefficient is significantly different from zero. Write a **brief 5:00** news summary of the resulting model. [Hint: the model corresponds to one of the pictures above. In your summary you may want to indicate what ad budget you would recommend and whether your recommendation depends on the size of the store.]

Regression Summary**Profit vs. 3 Independents**

Count	25
Num. Missing	0
R	.857
R Squared	.735
Adjusted R Squared	.697
RMS Residual	2.238

ANOVA Table**Profit vs. 3 Independents**

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	3	292.163	97.388	19.437	<.0001
Residual	21	105.217	5.010		
Total	24	397.379			

Regression Coefficients**Profit vs. 3 Independents**

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	16.068	2.364	16.068	6.798	<.0001
Ad Budget	2.466	.642	2.331	3.839	.0010
Ad Budget^2	-.143	.038	-2.359	-3.802	.0010
Size	.884	.197	.567	4.490	.0002

Confidence Intervals**Profit vs. 3 Independents**

	Coefficient	95% Lower	95% Upper
Intercept	16.068	11.153	20.984
Ad Budget	2.466	1.130	3.801
Ad Budget^2	-.143	-.221	-.065
Size	.884	.475	1.294

**Psych 5741 Final Exam: Answers
Fall 1995**

Question A

1.

A: $ID = \beta_0 + \varepsilon_i$

C: $ID = 3 + \varepsilon_i$

$H_0: \beta_0 = 3$

$PA - PC = 1, \quad n - PA = 299$

2.

A: $ID = \beta_0 + \beta_1 \text{GEN} + \beta_2 \text{AGE} + \beta_3 \text{FAM} + \varepsilon_i$

C: $ID = \beta_0 + \beta_1 \text{GEN} + \beta_2 \text{AGE} + \varepsilon_i$

$H_0: \beta_3 = 0$

$PA - PC = 1, \quad n - PA = 296$

3.

A: $ID = \beta_0 + \beta_1 \text{FR} + \varepsilon_i$

C: $ID = \beta_0 + \varepsilon_i$

$H_0: \beta_1 = 0$

$PA - PC = 1, \quad n - PA = 298$

4.

A: $ID = \beta_0 + \beta_1 \text{FR} + \beta_2 \text{GEN} + \beta_3 \text{FR GEN} + \varepsilon_i$

C: $ID = \beta_0 + \beta_1 \text{FR} + \beta_2 \text{GEN} + \varepsilon_i$

$H_0: \beta_3 = 0$

$PA - PC = 1, \quad n - PA = 296$

5.

A: $ID = \beta_0 + \beta_1 \text{ACT} + \beta_2 \text{ACT}^2 + \varepsilon_i$

C: $ID = \beta_0 + \beta_1 \text{ACT} + \varepsilon_i$

$H_0: \beta_2 = 0$

$PA - PC = 1, \quad n - PA = 297$

6.

Let $ACT4 = ACT - 4$

$$A: ID = \beta_0 + \beta_1 ACT4 + \beta_2 ACT4^2 + \varepsilon_i$$

$$C: ID = \beta_0 + \beta_2 ACT4^2 + \varepsilon_i$$

$$H_0: \beta_1 = 0$$

$$PA - PC = 1, \quad n - PA = 297$$

7.

$$A: ID = \beta_0 + \beta_1 ACT4 + \beta_2 ACT4^2 + \varepsilon_i$$

$$C: ID = 4.5 + \beta_1 ACT4 + \beta_2 ACT4^2 + \varepsilon_i$$

$$H_0: \beta_0 = 4.5$$

$$PA - PC = 1, \quad n - PA = 297$$

Question B

1. $P = \beta_0 + \beta_1 A + \varepsilon_i$

2. $P = \beta_0 + \beta_1 A + \beta_2 A^2 + \varepsilon_i$

3. $P = \beta_0 + \beta_1 A + \beta_2 S + \varepsilon_i$

4. $P = \beta_0 + \beta_1 A + \beta_2 S + \beta_3 A S + \varepsilon_i$

5. $P = \beta_0 + \beta_1 A + \beta_2 A^2 + \beta_3 S + \varepsilon_i$

6. $P = \beta_0 + \beta_1 A + \beta_2 A^2 + \beta_3 S + \beta_4 A S + \beta_5 A^2 S + \varepsilon_i$

[Common mistake: omitting the $A \cdot S$ term, a required component given $A^2 \cdot S$ is in the model.]

7.

$$P = \beta_0 + \beta_1 A + \beta_2 S + \beta_3 T + \beta_4 A S + \beta_5 A T + \beta_6 S T + \beta_7 A S T + \varepsilon_i$$

[Common mistake: using separate models for each graph rather than one model describing both graphs.]

8.

$$A: P = \beta_0 + \beta_1 A + \varepsilon_i$$

$$C: P = \beta_0 + \varepsilon_i$$

$$H_0: \beta_1 = 0$$

$$PA - PC = 1, \quad n - PA = 23$$

SSE(A) = 5236 from the problem and

$$SSE(C) = (n-1)\text{Var}(P) = 24(351) = 8,424.$$

Then SSR = 8424 - 5236 = 3188 and

$$\text{PRE} = \text{SSR}/\text{SSE}(C) = 3188/8424 = .378.$$

$$F^*(1,23) = [\text{PRE}/1]/[(1-\text{PRE})/23] = .378/((.622)/23) = 14.00.$$

From the table, $p < .01$.

Thus, we can reject MODEL C in favor of MODEL A and conclude that $b_1 = 2.5$ is reliably different from zero. As the Ad budget goes up by 1 unit (1000's of dollars), then Profits are predicted to increase by 2.5 units (1000's of dollars). $b_0 = 20$ [which we do not know is significantly different from zero from the above test] is the predicted Profit (in 1000s of dollars) for those stores not spending any money on advertising.

Summary: Each dollar spent on advertising returns approximately 2.5 dollars in profits.

9. MODEL: $\hat{P} = 16 + 2.5A - .14A^2 + .88S$

SLOPE: $2.5 - .28A$

MODEL Re-expressed as "simple" relationship between A and P:

$$\hat{P} = [16 + .28A^2 + .88S] + [2.5 - .28A]A$$

$b_0 = 16$ is the predicted profit when the ad budget equals 0 and the size of store is 0. [obviously meaningless and beyond the range of data!]

$b_1 = 2.5$ is the predicted change in Profits for a one unit change in Ad budget in the neighborhood of the Ad budget = 0 and when controlling for size of store. That is, it is the approximate slope for the first dollars spent within any level of store size.

$b_2 = -.14$ is 1/2 the magnitude of the change in the slope of the A:P relationship for each unit change in A when controlling for store size. That is, within any level of store size, subsequent dollars spent on advertising have less of a relationship to profits than the early dollars spent.

$b_3 = +.88$ is the predicted change in Profits for a one unit change in store Size, when controlling for ad budget and its quadratic relationship to profits. That is, within any level of ad budget, a unit (1000s of square feet) increase in store size predicts a .88 increase in Profits (1000s of dollars).

The beneficial effect of advertising plateaus when the slope = 0.

$2.5 - .28A = 0$ $A = 8.9$ would yield the greatest Profits. Spending more on advertising beyond that would be counter-productive by decreasing profits. This conclusion does not depend on the size of the store.

Summary: Early dollars spent on advertising are more effective than later dollars, regardless of store size. The optimal expenditure is 8.9; spending more than that increases profits by less than the cost of the advertising. Larger stores produce higher profits, on average, when stores large and small stores have the same ad budgets.