

Final Exam

PS 217, Spring 2002

1. Dr. Tori Ador was interested in studying the impact of cape color on instigating movement in bulls. She has a professional bullfighter use one of five cape colors (red, green, blue, black, and white). The DV is the time (in seconds) between waving the cape and the bull’s movement toward the bullfighter. Complete the analyses below and interpret the results as completely as you can, with careful consideration to the advice you’d give Dr. Ador based on these results. [15 pts]

ANOVA Table for Time to Charge

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Cape Color	4	.560	.140	.226	.9207	.903	.089
Residual	20	12.400	.620				

Means Table for Time to Charge

Effect: Cape Color

	Count	Mean	Std. Dev.	Std. Err.
Black	5	2.000	.707	.316
Blue	5	2.200	.837	.374
Green	5	2.000	.707	.316
Red	5	2.200	.837	.374
White	5	1.800	.837	.374

First of all, the results are not significant. Thus, the power of this study is quite low (~ .09). So, your conclusion would be that you would retain H_0 .

You should then begin to think about ways to make this study more powerful. Note that the times to initiate movement are all quite short and similar. One possible explanation for that result is that the bulls do not perceive the colors as different. Thus, you might consider using different colors or making these colors more vivid (in an effort to make the treatment effect larger). You might also think of ways to make the error term smaller, by using a more precise timing device, ensuring that the bulls are more similar in their speed, etc. And, of course, you should consider increasing the sample size from 5 to some larger n.

2a. Dr. Randy Mann is interested in the relationship between the number of sexual partners a person has had and the extent to which that person is satisfied with his or her life. He interviews a dozen people, asking about the number of sexual partners the person has had and assessing happiness using the Satisfaction in Ordinary Life Events scale (1-10, with 1 = Unhappy and 10 = Happy). The data are shown below. Analyze the data and interpret the results as completely as you can. [15 pts]

	Partners	SOLE	XY
	0	1	0
	2	8	16
	5	2	10
	1	3	3
	3	9	27
	2	6	12
	7	1	7
	1	3	3
	3	7	21
	2	8	16
	6	2	12
	4	5	20
Sum	36	55	147
Mean	3.0	4.583	
SS	50	94.9	

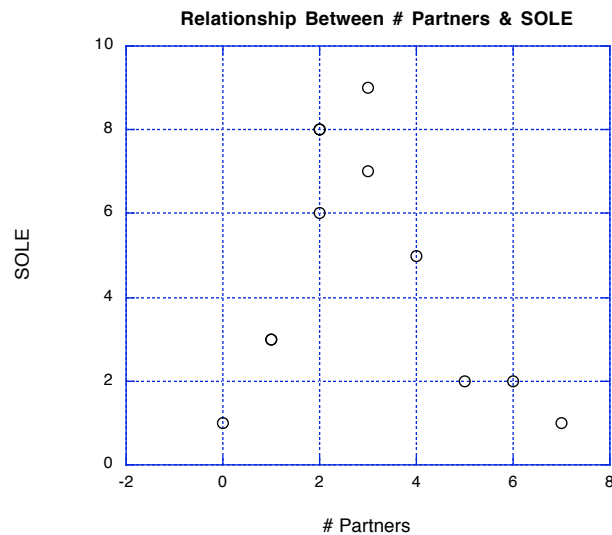
$H_0: \rho = 0$ $H_1: \rho \neq 0$ $r_{\text{crit}}(10) = .576$

Decision Rule: Reject H_0 if $|r| \geq r_{\text{crit}}$

$$r = \frac{147 - \frac{(36)(55)}{12}}{\sqrt{(50)(94.9)}} = \frac{18}{68.9} = .26$$

Decision: Retain H_0 , because $r < r_{\text{crit}}$. The linear relationship between # Partners and SOLE is not significant.

BUT, you should always graph your data, as seen below:



The relationship seems to be somewhat curvilinear, with SOLE increasing up to ~3 partners and then decreasing.

2b. Suppose that you convert every person's sexual partner score to a z score. Answer the following questions about the z scores. [5 pts]

The mean of the distribution of z scores would be: **0**

The standard deviation of the distribution of z scores would be: **1**

A person with 3 sexual partners (presumably not all at once) would have a z score of: **0**

A person with 6 sexual partners (don't even think about it) would have a z score of:

$$z = \frac{X - \bar{X}}{s} = \frac{6 - 3}{2.13} = 1.41$$

What percentage of people would have more than 3 sexual partners? **You can't really say, because you don't know if the distribution is normal.**

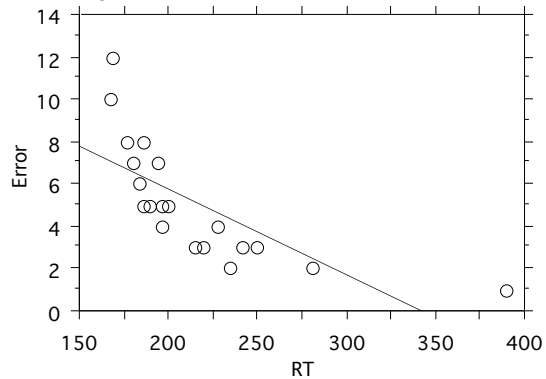
3. Researchers who measure reaction time (RT) for human participants often observe a relationship between the RTs and the number of errors that participants commit. This relationship is known as the *speed-accuracy trade-off*. In the data below, 20 participants were shown computer screens filled with the letter E. Some of the letters were in the usual orientation and some of the letters were backward. Participants are told to assess the number of backward E's as quickly as possible. The average number of backward letters missed over 10 trials constituted the mean number of errors. The average response time on each trial (in milliseconds) constituted the other variable. Interpret the results of the study as completely as you can. How many errors would you predict a person would make who responded with a RT of 200 milliseconds? How many errors would you predict a person would make who responded with a RT of 500 milliseconds? How much of the variability in errors is shared with RT? [10 pts]

Regression Summary

Error vs. RT

Count	20
Num. Missing	0
R	.724
R Squared	.524
Adjusted R Squared	.498
RMS Residual	2.007

Regression Plot



ANOVA Table

Error vs. RT

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	80.012	80.012	19.855	.0003
Residual	18	72.538	4.030		
Total	19	152.550			

Regression Coefficients

Error vs. RT

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	13.774	1.987	13.774	6.933	<.0001
RT	-.040	.009	-.724	-4.456	.0003

There is a significant negative linear relationship between Errors and RT, $r = -.724$, $p = .0003$. If a person responded in 200 ms, the number of predicted errors would be:

$$\hat{Y} = 13.774 - (.04)(200) = 5.77$$

If a person responded in 500 ms, you would not be able to predict the number of errors, because you did not observe anyone with a RT that large. The shared variability between these two variables would be $r^2 = .524$.

You might also note that the relationship appears to be curvilinear and that one point may be something of an outlier.

4a. Dr. Luke Attem was interested in factors that influence memory for faces. During the acquisition phase, participants were shown a series of 60 computer-generated male faces one at a time for 30 seconds each. Fifteen of the faces wore sunglasses, fifteen of the faces had full beards, fifteen of the faces wore hats, and fifteen of the faces were unadorned (no glasses, beard, hat, etc.). Each type of face occurred equally often within portions of the acquisition phase, to control for any position bias. (Note that the random ordering of the faces serves the same function as counterbalancing.) At test, 120 “unadorned” faces were presented, 60 new faces and the 60 original faces. That is, regardless of how the face was seen at acquisition, it was seen unadorned at test (no sunglasses, beard, or hat). The dependent variable was the percentage of faces of each type correctly recognized (100% indicating perfect recognition). Complete the analysis below and interpret the results of this experiment as completely as you can. [10 pts]

ANOVA Table for Acq Face

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	23	7041.490	306.152				
Category for Acq Face	3	5554.281	1851.427	263.688	<.0001	791.063	1.000
Category for Acq Face * Subject	69	484.469	7.021				

Means Table for Acq Face

Effect: Category for Acq Face

	Count	Mean	Std. Dev.	Std. Err.
Unadorned	24	60.167	9.867	2.014
Sunglasses	24	41.417	9.007	1.839
Beard	24	44.583	8.423	1.719
Hat	24	55.042	8.819	1.800

Decision: Reject H_0 , because $p < .05$ ($p < .0001$).

Post Hoc Test:

$$HSD = 3.7 \sqrt{\frac{7.0}{24}} = 2.0$$

Conclusion:

When people wore glasses, they were recognized significantly less often than if they wore beards or hats or if their faces were unadorned. When people wore beards, they were recognized significantly less often than when they wore hats or were unadorned. When people wore hats, they were recognized significantly less often than when their faces were unadorned.

4b. Given the data above, how likely is it that the sample of $n=24$ who saw unadorned faces at both acquisition and test were sampled from a population with $\mu = 65$? [10 pts]

$H_0: \mu = 65$ $H_1: \mu \neq 65$ $t_{\text{crit}}(23) = 2.069$

$$s_{\bar{x}} = \frac{9.867}{4.899} = 2.01$$

$$t = \frac{60.167 - 65}{2.01} = \frac{-4.83}{2.01} = -2.4$$

Decision: Reject H_0 , because $|t_{\text{obt}}| \geq t_{\text{crit}}$

Conclusion: The sample is likely to have come from a population with $\mu < 65$.

5. Individuals who are identified as having an antisocial personality disorder also tend to have reduced physiological responses to painful or anxiety-provoking stimuli. In everyday terms, these individuals show a limited physical response to fear, guilt, or anxiety. One way of measuring this response is with the galvanic skin response (GSR). With GSR, higher scores indicate lower responsivity and lower GSR scores indicate greater responsivity. In the study summarized below, three groups of individuals were tested: Normal Personality, Antisocial Personality, and Agoraphobics. First, briefly tell me why a group of Agoraphobics (or some other clinically diagnosed group) would be included in such a study:

To determine if the results are specific to antisocial people or to any clinical group.

As you can see, a third of each group is given the GSR under ordinary circumstances (baseline), a third is given a moderately stressful situation, and a third is given a highly stressful situation. Complete the source table and interpret the results of this study as completely as you can. [20 pts]

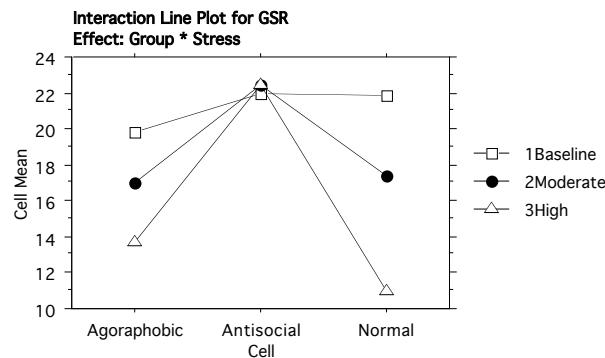
ANOVA Table for GSR

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Group	2	419.937	209.968	34.538	<.0001	69.076	1.000
Stress	2	323.556	161.778	26.611	<.0001	53.222	1.000
Group * Stress	4	226.825	56.706	9.328	<.0001	37.311	1.000
Residual	54	328.286	6.079				

Means Table for GSR

Effect: Group * Stress

	Count	Mean	Std. Dev.	Std. Err.
Agoraphobic, 1Baseline	7	19.857	2.116	.800
Agoraphobic, 2Moderate	7	17.000	1.633	.617
Agoraphobic, 3High	7	13.714	1.113	.421
Antisocial, 1Baseline	7	22.000	3.109	1.175
Antisocial, 2Moderate	7	22.429	2.637	.997
Antisocial, 3High	7	22.429	2.225	.841
Normal, 1Baseline	7	21.857	2.410	.911
Normal, 2Moderate	7	17.429	4.117	1.556
Normal, 3High	7	11.000	1.414	.535



First, compute $F_{Max} \cdot F_{Max} = 16.9 / 1.24 = 13.6$ $F_{Max Crit} = 17.5$, so there is no concern about heterogeneity of variance.

Next, compute HSD. $HSD = 4.58 \sqrt{\frac{6}{7}} = 4.24$.

Thus, Antisocial people show no difference between Baseline, Moderate, or High Stress. However, Agoraphobics and Normal people show a different pattern. Agoraphobics have higher GSR for Baseline than High Stress levels, but GSR levels for Baseline and Moderate Stress are equal and GSR levels for Medium and High Stress are equal. For Normal people, however, GSR levels are significantly higher for Baseline than Moderate and High Stress levels. GSR levels are also significantly higher for Moderate Stress than High Stress.

6. Do blondes really have more fun? To answer this intriguing question, Dr. Vanna Cream asks 20 people to rate how much fun they are having (on a scale of 1-10). Five of the participants have blonde hair, five have red hair, five have black hair, and five have brunet hair. Analyze the data below as completely as possible to address this question. If you are not blonde, would you dye your hair blonde based on the results of this study? Why or why not? [20 pts]

	Blonde	Red	Black	Brunet
	8	6	5	4
	9	5	4	5
	7	6	7	4
	8	7	6	5
	7	6	5	3
$\Sigma X (T)$	39	30	27	21
\bar{X}	7.8	6.0	5.4	4.2
SS	2.8	2.0	5.2	2.8

$$H_0: \mu_{\text{Blonde}} = \mu_{\text{Red}} = \mu_{\text{Black}} = \mu_{\text{Brunet}}$$

$$H_1: \text{Not } H_0$$

Source	SS	df	MS	F
Hair Color	33.75	3	11.25	14.1
Error	12.8	16	.8	
Total	46.55	19		

$F_{Max} = 5.2 / 2 = 2.6$ $F_{Max Crit} = 20.6$, so there is no concern that you've violated the homogeneity of variance assumption. Thus, you'd use $\alpha = .05$, so $F_{Crit} (3,16) = 3.24$

Decision: Reject H_0 , because $F_{Obt} \geq F_{Crit}$.

$$\text{Post Hoc Test: } HSD = 4.05 \sqrt{\frac{.8}{5}} = 1.62$$

Conclusion: Blondes have more fun than people with red or black hair. People with red hair have more fun than people with brunet hair. No other differences are significant. However, you should note that the hair colors are not randomly assigned, so you can't really make causal claims. That is, even though the blondes appear to have more fun than people with black hair, dying one's hair blonde may not increase one's level of fun.