Exam 3 PS 217, Fall 2008

1. OK, Zubin, this one's for you (the promised question). In an independent groups ANOVA, the best estimate of population variance (σ^2) is MS_{Within} (MS_{Error}). Is that also true for a repeated measures ANOVA? In other words, tell me whether or not MS_{Error} in a repeated measures design is a good estimate of population variance, along with your supporting logic. [5 pts]

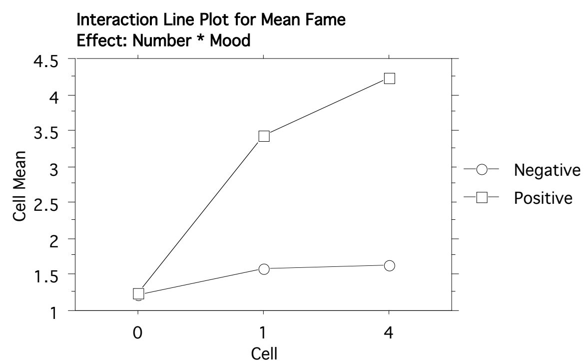
In an independent groups design, the MS_{Within} estimates variability due to individual differences and random variability. Those same two basic sources of variability underlie the population variance (σ^2). However, in a repeated measures design, the appropriate error term (MS_{Error}) reflects only random variability. As a result, it would not be an appropriate estimate of population variance (which is due to both random variability and individual differences).

2. Kitamura (2005) was interested in the impact of mood on cognitive processes. Kitamura thought that a positive mood leads to more automatic processing than a negative mood, which leads to more controlled processing. In one study, half of the participants were placed in a positive mood and half in a negative mood (using a mood induction technique). Then they were all given a list of non-famous companies either once or four times. Two days later they were asked to judge the fame of a list of companies, some of which were new (Number = 0) and some that had been seen previously (Number = 1 or 4). Let's pretend that the participants rated fame on a 7-point Likert-type scale (1 = "not famous" and 7 = " famous"). Suppose that the data had produced the results seen below. Complete the analysis and interpret the results as completely as you can. [15 pts]

Descriptive Statistics

Dependent Variable:Mean Fame

Mood	Number	Mean	Std. Deviation	N
Negative	0	1.20833	.178164	12
	1	1.58333	.327062	12
	4	1.63333	.486795	12
	Total	1.47500	.393791	36
Positive	0	1.23333	.182574	12
	1	3.42500	.748483	12
	4	4.23333	1.144420	12
	Total	2.96389	1.500124	36
Total	0	1.22083	.176879	24
	1	2.50417	1.097221	24
	4	2.93333	1.582147	24
	Total	2.21944	1.322038	72



Tests of Between-Subjects Effects

Dependent Variable:Mean Fame

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Mood	39.9	1	39.9	105	.000	.614	105.055	1.000
Number	38.1	2	19.1	50.3	.000	.603	100.337	1.000
Mood * Number	21.0	2	10.5	27.6	.000	.456	55.320	1.000
Error	25.1	66	.38					
Corrected Total	124.1	71						

You could avoid computing Hartley's F_{Max} by noting that all effects are significant at .01 or lower. If you computed, you'd find $F_{Max} = 41$. With $F_{Max\ Crit}$ about equal to 6, you would be concerned about violating the homogeneity of variance assumption, so you'd use $\alpha = .01$.

Thus, there is a significant main effect of Mood, a significant main effect of Number, and a significant interaction. (In each case, $p < .001$.)

To interpret the interaction displayed above, you'd compute Tukey's HSD:

$$HSD = 4.16\sqrt{\frac{.38}{12}} = .74$$

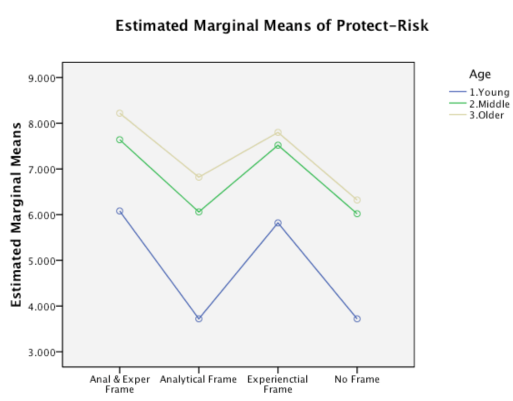
The judged fame of new companies (0, or not seen previously) was the same whether the participant was in a positive ($M = 1.23$) or a negative mood ($M = 1.21$). However, if participants saw the company name one time, the company was judged to be more famous if people were in a positive mood ($M = 3.4$) than if people were in a negative mood ($M = 1.58$). The same was true if participants saw the company name four times. They judged the company to be more famous if people were in a positive mood ($M = 4.23$) than if people were in a negative mood ($M = 1.63$).

3. One area of psychology looks at factors that influence decision-making. One factor that people have studied is how a decision is influenced by the way in which the information is delivered. Even though the information is identical, people's decisions will differ when the information is placed in a different context (frame). Suppose that a researcher was interested in looking at the impact of four different frames on people's willingness to engage in risky behavior (or to be more protective). One scenario involves the participant's willingness to smoke cigarettes. The four frames are: NF = No Frame (so it just asks the participant to imagine that he or she has been smoking for a while and enjoys doing so), AF = Analytical Frame (with statistical information about the scenario, such as how many people die of lung cancer each year), EF = Experiential Frame (which attempts to make the scenario personally relevant by asking the participant to think about a family member dying from lung cancer), and AEF = Analytical + Experiential Frames (which puts the two types of information together). Participants read a series of scenarios and then gave a response that indicated their willingness to engage in risky behavior. The dependent variable is called Protect-Risk, where a positive score indicates a more protective response and a negative score represents a willingness to engage in riskier behavior. Suppose that the researcher is also interested in looking at the impact of age (Young 18-23, Middle 38-43, and Older 58-63). Complete the source table below and interpret the results as completely as you can. [15 pts]

Descriptive Statistics

Dependent Variable:Protect-Risk

Frame	Age	Mean	Std. Deviation	N
Anal & Exper Frame	1.Young	6.08000	.356371	5
	2.Middle	7.64000	.296648	5
	3.Older	8.22000	.277489	5
	Total	7.31333	.978969	15
Analytical Frame	1.Young	3.72000	.370135	5
	2.Middle	6.06000	.680441	5
	3.Older	6.82000	.402492	5
	Total	5.53333	1.443046	15
Experienctial Frame	1.Young	5.82000	.653452	5
	2.Middle	7.52000	.311448	5
	3.Older	7.80000	.400000	5
	Total	7.04667	1.007732	15
No Frame	1.Young	3.72000	.486826	5
	2.Middle	6.02000	1.023230	5
	3.Older	6.32000	.426615	5
	Total	5.35333	1.365319	15
Total	1.Young	4.83500	1.230009	20
	2.Middle	6.81000	.991490	20
	3.Older	7.29000	.850944	20
	Total	6.31167	1.478099	60



Tests of Between-Subjects Effects

Dependent Variable:Protect-Risk

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Frame	46	3	15.339	57.4	.000	.782	172.346	1.000
Age	67.7	2	33.860	126.8	.000	.841	253.634	1.000
Frame * Age	2.4	6	.392	1.47	.210	.155	8.798	.515
Error	12.8	48	.267					
Corrected Total	128.9	59						

No need for Hartley's F_{Max}, because both main effects are significant with $p < .01$. However, in this case $F_{Max} = 13.64$ and $F_{Max\ Crit} = 51.2$, so there would be no concern about violating the homogeneity of variance assumption.

FRAME

$$HSD = 3.78\sqrt{\frac{.267}{15}} = .5$$

	AE	A	E	N
AE	--			
A	1.77	--		
E	.26	1.5	--	
N	1.96	.2	1.7	--

Thus, AE and E lead to more risk-averse responses than A or N. It appears that offering a scenario that has an experiential component leads people to avoid risk.

AGE

$$HSD = 3.43\sqrt{\frac{.267}{20}} = .4$$

	Y	M	O
Y	--		
M	1.97	--	
O	2.45	.48	--

Older people are more risk averse than middle-aged people or young people. And middle-aged people are more risk averse than young people.

4. Two researchers were interested in studying the effects of reward magnitude on performance. Both researchers used introductory psychology students as participants, the same total number of participants (21), the same type of reward and reward magnitudes ($1, $5, $20), the same apparatus, the same task, and the same performance measure (DV). One researcher used an independent groups design and, on the basis of the results, cannot reject the null hypothesis (that reward has no effect on performance). The other researcher used a repeated measures design and found a statistically significant effect of reward magnitude — larger rewards lead to better performance. Assume that neither study has a major flaw (e.g., repeated measures design is properly counterbalanced, random assignment to conditions). There are two fundamental reasons why the two researchers might have reached different conclusions. One reason concerns the sensitivity of the test of the null hypothesis. The other reason concerns the nature of the participant's experience in the two studies. Provide me with a clear explanation of the two reasons for the different results that the two researchers obtained. Would you trust the results of one study more than the other? Why? Finally, complete the source tables for the two experimenters seen below. [10 pts]

Independent Groups Design (F_{Crit} = 3.55):

Source	*SS*	*df*	*MS*	*F*
Treatment	28	2	14	3.5
Error	72	18	4	
Total	100	20		

Repeated Measures Design (F_{Crit} = 3.23):

Source	*SS*	*df*	*MS*	*F*
Treatment	20	2	10	5
Within	180	60		
Subject	100	20		
Error (Subj x Treat)	80	40	2	
Total	200	62		

In general, because the repeated measures design is more powerful than the independent groups design, you'd expect that the repeated measures design would lead to a larger *F*, which is the case here.

In this particular case, however, the repeated measures design may introduce a problem. That is, each participant will receive all rewards ($1, $5, $20) in some order. After receiving $20, a participant might next be asked to complete the same basic task, but now for a reward of only $1. That might annoy the participant, but it would certainly alert the participant to the fact that the researcher is interested in amount of reward. In an independent groups design, because each participant receives only one level of the factor, he or she would have no idea that other participants are going to receive different rewards. (Another problem with the repeated measures design is that with 21 participants, you wouldn't be able to appropriately counterbalance. You'd need a multiple of 6.)

In this case, I'd likely prefer the independent groups design, but would increase the sample size to garner more power. Doing so would likely increase the *F* from 3.5.

5a. First of all, imagine a repeated measures design with seven levels. Can you tell me *why* you'd need to counterbalance such a design, what kind of counterbalancing you'd use, and how many participants you'd need? What is the impact of counterbalancing on order and carry-over effects? [3 pts]

You'd need to counterbalance because of a concern about carry-over or order effects in a repeated measures design. In this case, you'd want to use incomplete counterbalancing, which would lead to a total of 14 orders. Thus, you'd need to run a multiple of 14 participants. Keep in mind that counterbalancing doesn't eliminate the carry-over or order effects, it merely distributes them equally over the conditions.

5b. OK, now let's assume that there is a particular order effect—a practice effect. That means that scores on the DV will improve over time as a result of practice. What is the impact on your error term (MS_{Error}) of counterbalancing? [2 pts]

Generally speaking, counterbalancing will increase your MSError compared to no counterbalancing. Of course, you need to counterbalance with a repeated measures design. It's just that doing so will spread the order effects over the conditions, which will increase the variability in the denominator.

6. Dr. Ginger Vitas is a health psychologist who is interested in the relationship between dental heath (operationally defined as number of cavities found in an annual checkup) and general health (operationally defined as the number of illnesses experienced in the preceding year). Analyze the output seen below as completely as you can. If a person had 3 cavities in a year, how many illnesses would you predict? If a person had 7 cavities, how many illnesses would you predict? What proportion of variability do these two variables share? [10 pts]

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.861[a]	.742	.736	1.25312

a. Predictors: (Constant), cavities

ANOVA[b]

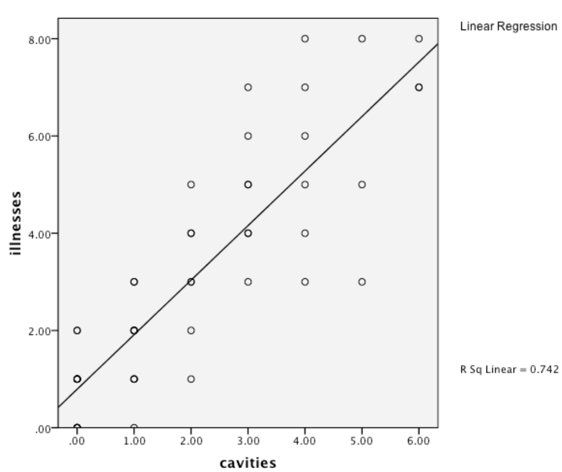
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	216.306	1	216.306	137.748	.000[a]
	Residual	75.374	48	1.570		
	Total	291.680	49			

a. Predictors: (Constant), cavities
b. Dependent Variable: illnesses

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.792	.263		3.005	.004
	cavities	1.122	.096	.861	11.737	.000

a. Dependent Variable: illnesses



There is a significant positive linear relationship (*r* = .861) between illnesses and cavities, because p < .001.

The regression equation is: Y = 1.122X + .792.

Thus, with X = 3, your best prediction of Y is 4.2

With X = 7, you could predict that Y is 8.6 if the trend continues. Or, you could say that you couldn't predict because you didn't observe anyone with 7 cavities.

The proportion of shared variability (r^2) is .742.

7a. [Dunn] A health psychologist is interested in the relationship between self-reported stress and objective health ratings. Higher scores on each measure indicate greater amounts of that variable (i.e., more stress, better health). Using the following data, determine whether or not this relationship is significant. If a person has a stress score of 5, what is your best estimate of that person's health score? [15 pts]

	Stress	Health	XY (S*H)
	2	10	20
	9	3	27
	10	4	40
	11	5	55
	4	13	52
	5	11	55
	1	10	10
	3	4	12
Sum	45	60	271
SS	103.88	106	

H_0: ρ = 0 **H_1: ρ ≠ 0** **$r_{Crit}(6) = .707$**

$$r = \frac{271 - \frac{(45)(10)}{8}}{\sqrt{(103.88)(106)}} = \frac{-66.5}{104.9} = -.63$$

Decision: Retain H_0, no significant linear relationship

Your best guess of the health score is the mean health score (7.5).

7b. How likely is it that the sample of stress scores in this problem came from a population with a mean (μ) of 4.5? [5 pts]

H_0: μ = 4.5
H_1: μ ≠ 4.5

$t_{Crit}(7) = 2.36$

$$t = \frac{5.625 - 4.5}{3.85/\sqrt{8}} = \frac{1.125}{1.36} = .827$$

Decision: Retain H_0, because $t_{Obt} < t_{Crit}$. The stress scores may have come from a population with μ = 4.5.

7c. OK (*z*-score questions for Janet), if you were to convert each of the stress scores to *z*-scores, what would be the mean of those *z*-scores? [1 pt]

0, because the mean of a set of *z*-scores is always 0.

7d. What would be the standard deviation of the *z*-scores? [1 pt]

1, because the standard deviation of a set of *z*-scores is always 1.

7e. Once converted to *z*-scores, what proportion of stress scores would be above a *z*-score of 0? {Think carefully!} [1 pt]

You don't know that the distribution is normal, so you can't be certain about the distribution of *z*-scores. Only if the distribution is normal would you know that the proportion of scores above 0 would be .50.

7f. After converting all of your stress scores to z-scores, suppose that you then added a constant of 5 to each *z*-score. What is the mean of this new distribution of scores? [1 pt]

5, because adding a constant to each score increases the mean by that amount. (And that's the way they get standardized scores such as the SAT, with a mean of 500.)

7g. After converting all of your stress scores to *z*-scores, suppose that you then added a constant of 5 to each *z*-score. What is the standard deviation of this new distribution of scores? [1 pt]

1, because adding a constant to each score has no effect on variability.