

1. Dr. Noah Weisser was interested in investigating the effects of sleep deprivation and alcohol on driving ability. To that end, he conducted a 3x3 independent groups study. Participants were given identical 32 oz. drinks that could contain 2, 4, or 8 ozs. of vodka (with the remainder of the glass filled with orange juice). Within each level of alcohol, one-third of the participants were deprived of sleep for 2 days, one-third were deprived of sleep for 3 days, and the final third were deprived of sleep for 4 days. Each of the participants operated a driving simulator for 30 minutes, during which the number of driving errors (objects hit, crossing into the oncoming lane, etc.) were recorded. Complete the analysis shown below and interpret the results as completely as you can. [20 pts]

ANOVA Table for Errors

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Alcohol	2	688.5	344.264	167.9	<.0001	335.932	1.000
Sleep Dep	2	461.8	230.889	112.6	<.0001	225.301	1.000
Alcohol * Sleep Dep	4	93.2	23.306	11.4	<.0001	45.483	1.000
Residual	63	129.2	2.050				

Means Table for Errors**Effect: Alcohol * Sleep Dep**

	Count	Mean	Std. Dev.	Std. Err.
2 oz., 2 days	8	1.375	1.061	.375
2 oz., 3 days	8	3.250	1.035	.366
2 oz., 4 days	8	6.375	1.188	.420
4 oz., 2 days	8	3.500	.926	.327
4 oz., 3 days	8	6.875	.991	.350
4 oz., 4 days	8	7.750	.707	.250
8 oz., 2 days	8	7.250	1.282	.453
8 oz., 3 days	8	9.500	1.195	.423
8 oz., 4 days	8	16.500	3.071	1.086

The interaction is significant, so that's where you should focus your attention. A graph would probably help your interpretation, but I'll save space and omit the graph for now. You'd need to compute the HSD to determine which means differ:

$$HSD = 4.55 \sqrt{\frac{2.05}{8}} = 2.3$$

Thus, at 2 ozs. 4 days of sleep deprivation leads to more errors than 3 or 2 days, neither of which differ. At 4 ozs. a different pattern emerges, where the errors at 3 and 4 days of deprivation produce equivalent errors that both exceed the number of errors at 2 days of deprivation. The pattern of results at 8 ozs. is similar to the pattern at 2 ozs. of alcohol, with 4 days of deprivation leading to significantly more errors than 2 or 3 days of deprivation, which do not differ (barely).

$F_{Max} = 9.43 / .5 = 18.9$ $F_{Max Crit} = 13.5$ According to this test, we should be concerned that we may have violated the homogeneity of variance assumption. However, even using $\alpha = .01$, all of the effects would remain significant.

2a. Two researchers were interested in studying the effects of reward magnitude on performance. Both researchers draw from the same pool of participants, use the same total number of participants (24), 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 uses 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 uses a repeated measures design and finds 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). There are three fundamental reasons why the two researchers might have reached different conclusions. One reason concerns the sensitivity of the test of the null hypothesis. Another reason concerns the nature of the participant’s experience in the two studies. A final reason has to do with the tentativeness of hypothesis testing, regardless of whether or not a researcher rejects the null hypothesis. Provide me with a clear explanation of the reasons that the two researchers may have reached such different conclusions. Would you trust the results of one study more than the other? Why? [15 pts]

1. A repeated measures design is more powerful than an independent groups design, so you should anticipate that the repeated measures design is more likely to lead you to reject H_0 . Note that you would get 72 pieces of data from the repeated measures design and only 24 pieces of data from the independent groups design.

2. In this particular case, the repeated measures design might not be all that appropriate. Because participants would do the task, get a reward, do the task, get a reward, etc., they are likely to exhibit carry-over effects (“Why are you paying me only \$1 this time, you just paid me \$20?). With proper counterbalancing, the effects would be distributed equally over the three conditions, but the effects may be somewhat permanent. Thus, the demand characteristics of the study may lead the participants to figure out what’s going on (the reward is important) and then behave accordingly.

3. We could always be making a Type I error with the repeated measures design and correctly retaining H_0 with the independent groups design. We could also be making a Type II error with the independent groups design and correctly rejecting H_0 with the repeated measures design.

2b. Finally, complete the source tables for the two experiments, as seen below. {Remember, the RM design is more efficient, so participants generate more than one piece of data.} [10 pts]

Independent Groups Design:

Source	SS	df	MS	F	
Between	24	2	12	3.4	$F_{\text{Crit}} = 3.47$
Error	74	21	3.52		
Total	98	23			

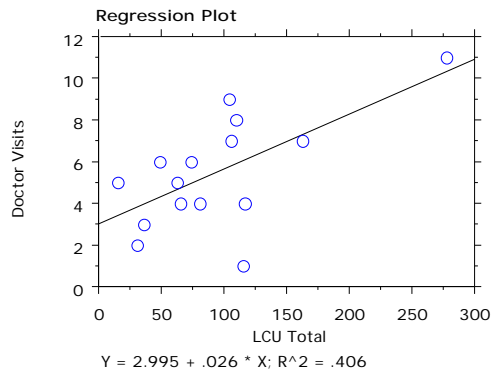
Repeated Measures Design:

Source	SS	df	MS	F	
Treatment	14	2	7	3.5	$F_{\text{Crit}} = 3.20$
Subjects	88	23	3.83		
Error	92	46	2		
Total	194	71			

3. Studies have suggested that the stress of major life changes is related to subsequent physical illness. Holmes and Rahe (1967) devised the Social Readjustment Rating Scale (SRRS) to measure the amount of stressful change in one's life. Each event is assigned a point value, which measures its severity. For example, at the top of the list, death of a spouse is assigned 100 life change units (LCU). Divorce is 73 LCUs, retirement is 45, change of career is 36, the beginning or end of school is 26, and so on. The more life change units one has accumulated in the past year, the more likely he or she is to have an illness. The following StatView analyses show the results from a hypothetical set of data. Interpret these results as completely as you can. For these data, if a person had accumulated 100 LCUs, how many doctor visits would you predict? If a person had accumulated 400 LCUs, how many doctor visits would you predict? [10 pts]

Regression Summary
Doctor Visits vs. LCU Total

Count	15
Num. Missing	0
R	.637
R Squared	.406
Adjusted R Squared	.360
RMS Residual	2.135



ANOVA Table
Doctor Visits vs. LCU Total

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	40.469	40.469	8.877	.0107
Residual	13	59.264	4.559		
Total	14	99.733			

Regression Coefficients
Doctor Visits vs. LCU Total

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	2.995	.996	2.995	3.006	.0101
LCU Total	.026	.009	.637	2.979	.0107

There is a significant positive linear relationship between Doctor Visits and LCU Total, $r = .637$ ($F(1,13) = 8.9$, $MSe = 4.6$, $p < .05$). Because $r^2 = .41$, these two variables share a fair amount of variability (~40%). Thus, about 60% of the variability in Doctor Visits is associated with variables other than LCU. If a person had accumulated 100 LCUs, then the regression equation would predict 5.6 (~6) doctor visits. If a person were unfortunate enough to have accumulated 400 LCUs, we would not be able to make a good prediction, because we haven't observed anyone with that large a LCU score. If the current trend continued, we could predict 13.4 doctor visits...but I'd be very tentative in making such a prediction. As it stands, I think that we should be a bit cautious with these data, because one of the points appears to be an outlier (LCU of about 280).

4. OK, here's the actual data from the last lab exercise on the impact of duration on face perception. Analyze and interpret the data as completely as you can. Tell me what you would conclude from these results. (I was expecting more than 27 participants, but that's all I could find on the Zip disk.) The DV is percent hits. That is, the number of the 16 faces at each duration that were correctly identified as old divided by 16. What would you say about the results given that the average proportion of false positives was .32. (False positives occur when a new face is called old.) {Can you remember how we used false positives in an earlier lab exercise?} [15 pts]

ANOVA Table for Duration

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	26	1.162	.045				
Category for Duration	2	.611	.305	26.796	<.0001	53.591	1.000
Category for Duration * Subject	52	.592	.011				

Means Table for Duration

Effect: Category for Duration

	Count	Mean	Std. Dev.	Std. Err.
H1.PERC	27	.500	.165	.032
H2.PERC	27	.641	.169	.032
H3.PERC	27	.708	.108	.021

There was a significant effect of duration on memory ($F(2,52) = 26.8$, $MSe = .011$, $p < .05$). (In other words, we would reject $H_0: \mu_{Short} = \mu_{Medium} = \mu_{Long}$.) The next step would be to compute HSD to conduct post hoc analyses. Using $q = 3.42$, $HSD = .07$. Thus, I would conclude that participants remembered the faces better when they were presented for the Long duration compared to the Short duration. They also remembered the faces better at the Medium duration than the Short duration. The Medium and Long exposures didn't produce differences in memory. These memory scores are likely to be inflated, however, because of the high False Positive rate. That is, with a False Positive rate of .32, it appears that participants were saying "old" to lots of faces...even faces that they hadn't seen before.

5. A health educator suspects that the "days of discomfort" caused by common colds can be reduced by ingesting large doses of Vitamin C and visiting a sauna every day. Participants with new colds are randomly assigned to one of four different doses of Vitamin C (500, 1000, 1500, or 2000 milligrams) and to one of three different daily exposures to a sauna (0, .5, or 1 hour). The DV is the number of days of discomfort experienced by each of the participants. Complete the source table below and analyze and interpret the results of this study as completely as you can. Then tell me what your next step would be. [10 pts]

ANOVA Table for Days of Discomfort

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Dose of C	3	4.983	1.661	2.492	.0713	7.475	.576
Sauna	2	1.200	.600	.900	.4133	1.800	.190
Dose of C * Sauna	6	3.067	.511	.767	.5998	4.600	.270
Residual	48	32.000	.667				

Means Table for Days of Discomfort
Effect: Dose of C * Sauna

	Count	Mean	Std. Dev.	Std. Err.
1000 mg, .5 Hr	5	3.600	1.140	.510
1000 mg, 0 Hr	5	3.800	.837	.374
1000 mg, 1 Hr	5	3.600	1.140	.510
1500 mg, .5 Hr	5	3.800	.837	.374
1500 mg, 0 Hr	5	3.200	.837	.374
1500 mg, 1 Hr	5	4.200	.837	.374
2000 mg, .5 Hr	5	3.000	.707	.316
2000 mg, 0 Hr	5	3.600	.548	.245
2000 mg, 1 Hr	5	3.600	.548	.245
500 mg, .5 Hr	5	4.200	.837	.374
500 mg, 0 Hr	5	4.000	.707	.316
500 mg, 1 Hr	5	4.400	.548	.245

This study resulted in no significant results, so you would need to address reasons for that outcome (lack of power). First of all, you should note the lack of an appropriate control group (0 mg of Vitamin C). Anything else that you might do to produce a larger difference in the treatment effects should also be considered (could you give a larger dosage of C? could you have people in the sauna for longer than 1 hour?). It should be obvious to you that the sample size (n = 5) is quite small, so increasing sample size would be one area to consider when looking for greater power. Finally, you should look at ways that you might reduce variability among your participants. Is there any way to ensure that the level of illness is equated across all participants? Would a more sensitive measure help (hours of discomfort rather than days of discomfort)?

6. Many research studies have found a negative relationship between average IQ and family size. Zajonc and Markus (1975) have developed a mathematical model for predicting the intellectual climate of a family based on the number of family members and their ages. For this exercise, see if you can develop a simple linear model for predicting intellectual climate (Y) from the number of children in a family (X), using the data below: [10 pts]

	Number of Children	Intellectual Climate	XY
	3	8	24
	0	18	0
	4	8	32
	2	14	28
	1	12	12
Sum	10	60	96
Mean	2	12	
Variance	2.5	18	
SS	10	72	

Because the question asks about a relationship, you need to compute r.

$$r = \frac{SP}{\sqrt{SS_X \cdot SS_Y}} = \frac{96 - \frac{(10)(60)}{5}}{\sqrt{(10)(72)}} = \frac{-24}{26.8} = -.90$$

Because $r_{\text{crit}}(3) = .878$, I would reject H_0 . Thus, it would make sense to construct the regression equation:

$$\hat{Y} = -2.4X + 16.8$$

Thus, you could make predictions about intellectual climate from number of children using the equation above.