Please show **all** your work! Answers without supporting work will not be given full credit. Write answers in spaces provided. You have 1 hour and 20 minutes to complete this exam. You may use a personal calculator, any tables provided, and any reference sheets provided. By placing your name on the line below, you agree to uphold the Honor Code during this test.

Name: Key

1. A die is rolled many times to see if it is loaded. It lands on each side in the following way.

Side	1	2	3	4	5	6
Rolls	5	6	6	10	8	7

Create the expected counts table which would be used to test the claim that the die is not loaded, i.e. that the distribution of rolls should be uniform. (10pts)

Side	1	2	3	4	5	6
Rolls	7	7	7	7	7	7

← There are 42 observations and $P(1)=P(2)=\ldots=P(6)=\frac{1}{6}$.

2. True or false: A one-way ANOVA tests that the population ~~variances~~ means are the same for three or more populations. (5pts)

Answer: False

3. A two-way table test is conducted on a 3×4 table, and the test statistic is found to be 11. What can we conclude about the row and column variables? (5pts)

5%, $df=6$, χ^2-dist: critical value = 12.592
10%, $df=6$, χ^2-dist: critical value = 10.645

Answer: If we use $\alpha=0.05$, they would be indep. If we use $\alpha=0.10$, they would not be indep.

4. We have the following two-way table. Does the table meet the requirements for a two-way table test? Why or why not? (5pts)

	Red	Green	Yellow	Total
Car	10	4	4	18
Truck	1	1	4	6
SUV	1	1	0	2
Total	12	6	8	26

Answer: No! The average of the expected counts would be $2.\overline{88}$, which is less than 5!!

5. The line of best fit for 25 (x, y) pairs of data is given by $\hat{y} = 10 + 2x$. We also know that $\bar{x} = 12$ and $\bar{y} = 8$. We preform a 5% test of significance for ρ and get a P-value of 0.045. Our value of r was 0.97. What proportion of the variance in y can be explained the the variance in the linear regression model? (5pts)

Answer: It's r^2! So it's 94.09%!

6. We have the following table for a test for significance for β_1 where the alternative hypothesis is that $\beta_1 \neq 0$. The critical F value for this test is 11.26. Fill in the missing parts of the table. Does this test support the claim that $\beta_1 = 0$? (5pts)

Source	DF	SS	MS	F
Model	1	77.092	77.092	0.391625
Error	8	3.133		
Total	9	80.225		

← Thus, we fail to reject $H_0: \beta_1 = 0$ and support our claim

Answer: Yes!

7. Is a prediction interval for a future observation or a confidence interval for the mean response wider? (5pts)

Answer: prediction interval is wider

8. Which two tests are numerically equivalent? Circle only one. (5pts)

- a goodness-of-fit test and a two-way table test
- a test for population linear correlation and a test for slope of a population regression line
- a two-way ANOVA an two one-way ANOVAs
- all of these
- none of these

Cont.

9. For the data below, we have that $r = 0.6728$. Does the data support the claim that x and y are linearly correlated? Use $\alpha = 0.01$ and a critical value test. (15pts)

x	1	2	3	4	5	6	7	8	9	10
y	2	4	5	12	6	3	8	10	11	10

Test statistic: $t = \dfrac{0.6728\sqrt{10-2}}{\sqrt{1-0.6728^2}} = 3.477$

Critical value: (from t distribution with 8 degrees of freedom)

3.355.

this lines up w/ our claim

Thus! We reject $H_0: \rho = 0$ and support $H_a: \rho \neq 0$.

Answer: Yes!!

10. A medical researcher wants to tell if a test for diabetes is effective. She collects the following data from a sample of 100 patients.

	Diabetic	Not Diabetic
Test Positive	35	15
Test Negative	15	35

← observed

Use $\alpha = 0.0005$ to test the claim that the result of the test is independent of whether or not the patient has diabetes. Use a critical value test. (15pts)

expected →

	D	ND
TP	25	25
TN	25	25

$\Rightarrow \chi^2 = \frac{(35-25)^2}{25} + \frac{(15-25)^2}{25} + \frac{(35-25)^2}{25} + \frac{(15-25)^2}{25} = 16$

The critical value (from χ^2 table with 1 degree of freedom) is 12.12.
Thus, reject H_0: the factors are independent. Support H_a: the factors are dependent.

Answer: The data does not support the claim. It looks like the result of the test for diabetes does depend on the status of the patient.

Cont.

11. If $\hat{y} = 10 + x$ an $SE_{\hat{y}} = 0.5$, produce a 90% prediction interval for $x = 11$, assuming that 11 is in the range of observed x values and that the data is actually linearly correlated. (10pts) Assume $n=12$.

$\hat{y} = 21$, $t^* = 2.228$, $SE_{\hat{y}} = 0.5$, so

$PI = (21 - 2.228(0.5), 21 + 2.228(0.5))$

$= (19.886, 22.114)$

Answer:

12. Fill in the missing parts of the two-way ANOVA test. At a 5% significance level, does it appear that there is an effect due to the interaction of the two factors? (The critical value for this test is 2.15 at the 5% significance level.) Also, tell what type of distribution the test statistic has — include degrees of freedom. (15pts)

Source	DF	SS	MS	F
A	2	53.733333	26.8666	15.3038
B	4	146.83333	36.70833	~~15.3038~~ 20.9098
AB	8	135.766	16.970833	9.6669
Error	45	79	1.75555	
Total	59	415.33333		

The test statistic has an $F(8,45)$ distribution.

So we reject H_0: there is no effect due to the interaction of factors A & B.

Answer: Yes, it does appear that there is an effect due to interaction!

13. Which is larger $SE_{\hat{\mu}}$ or $SE_{\hat{y}}$? (1pt)

The End.