STAT 515 sp 2024 Final Exam

Karl Gregory

- Do not open this exam until told to do so.
- You may have two handwritten sheets of notes out during the exam.
- You have 150 minutes to work on this exam.
- You may NOT use any kind of calculator.
- If you are unsure of what a question is asking for, do not hesitate to ask me for clarification.
- *Good luck, and may the odds be ever in your favor!*

$X \sim$		$\mathcal{X}$	$\mathbb{E}X$	$\text{Var}(X)$
Binomial(n, p)	$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$	$x = 0, 1, \ldots, n$	np	$np(1-p)$
Poisson(λ)	$P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$	$x = 0, 1, 2, \ldots$	λ	λ
Exponential(λ)	$P(X \leq x) = 1 - e^{-x\lambda}$	$x > 0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$

$$\hat{p}_n \pm z_{\alpha/2} \cdot \sqrt{\hat{p}_n(1-\hat{p}_n)/n} \qquad \bar{X}_n \pm t_{n-1,\alpha/2} \cdot S_n/\sqrt{n}$$

$$Z_{\text{test}} = \frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} \qquad T_{\text{test}} = \frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}$$

A t-table is attached to this exam.

1. Eighty percent of purchasers of a 14” carbon steel wok from KG’s Discount Store are satisfied with the wok. Ten percent of the purchasers post a review of the wok on the store’s website. Of the reviews, ninety percent indicate satisfaction with the wok. Give the probability that:

   (a) A satisfied purchaser posts a review of the wok.

   (b) A customer who does not post a review is satisfied with the wok.

2. Ten people order a 14" carbon steel wok from KG's Discount Store. Remember that 80% of purchasers are satisfied with the wok. Assuming each purchaser's satisfaction or dissatisfaction with the wok to be independent of that of the other purchasers, give an expression (you do not have to evaluate it) for the probability that:

   (a) Exactly 8 purchasers are satisfied with the wok.

   (b) All 10 purchasers are satisfied with the wok.

   (c) At least one purchaser is dissatisfied with the wok.

3. In order to season the 14" carbon steel wok for first-time use, purchasers must spend several minutes super-heating it until the metal acquires a bluish tint. Nine randomly selected purchasers of the wok recorded in a survey the number of minutes they spent super-heating their woks. The mean and standard deviation of the reported numbers of minutes were $\bar{X}_n = 35$ and $S_n = 5$. Mr. KG of KG's Discount Store holds firmly to the conviction that properly seasoning a wok requires at least 40 minutes of super-heating. If he concludes that purchasers of the wok do not spend, on average, sufficient time super-heating their woks, he will begin shipping with each wok a comprehensive wok-seasoning guide prescribing no less than forty minutes of super-heating during the seasoning process.

   (a) Give the null and alternate hypotheses which are of interest to Mr. KG.

   (b) Compute the test statistic of the test for testing the hypotheses in part (a).

   (c) Select the interval in which the p-value lies (this is multiple choice):

      i. $(0.0005, 0.001]$
      ii. $(0.001, 0.005]$
      iii. $(0.005, 0.01]$
      iv. $(0.01, 0.025]$
      v. $(0.025, 0.05]$
      vi. $(0.05, 0.10]$

   (d) Should Mr. KG begin shipping the wok seasoning guide with each wok? Explain your answer.

   (e) What assumption, if any, is implicit in the analysis you have carried out?

4. A 14" cast-iron wok is also sold by KG's Discount Store and is marketed as an alternative to the 14" carbon steel wok. To better understand customer sentiment around these products, Mr. KG acquires, for each product, 1-to-5-star ratings from a random sample of 40 purchasers. The frequencies of each rating for each product, along with the mean and standard deviation of the ratings for each product, are tabulated here:

Rating	1	2	3	4	5	mean	std. dev
Carbon steel	2	6	18	12	2	3.150	0.921
Cast-iron	4	14	17	3	2	2.625	0.952

(a) Are the ratings of the sampled purchasers drawn from a Normal distribution? Explain why or why not.

(b) To what phenomenon owes the fact that difference in mean ratings will be approximately Normally distributed?

(c) Study the R output below and write an assessment to Mr. KG, based on these data, of the customer sentiments around the 14" cast-iron and carbon steel woks. Give a justification of your assessment.

```
             Two Sample t-test

data:  carbon_steel and cast_iron
t = 2.5059, df = 78, p-value = 0.0143
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1078986 0.9421014
sample estimates:
mean of x mean of y
    3.150     2.625
```

5. Ever thoughtful of his customers, Mr. KG wishes to send a recipe with each 14" carbon steel wok he ships. To decide which recipe among three possible recipes he should send, he recruits twelve individuals who are culinarily inclined, but who have as yet never cooked with a carbon steel wok; he assigns each individual randomly to one of the three recipes, such that four individuals are assigned to each recipe. Each individual then cooks the assigned recipe in Mr. KG's kitchen with Mr. KG's very own carbon steel wok under the kind tutelage of that same Mr. KG—and afterwards rates the level of overall reward and gratification experienced on a scale of 1 to 10. The ratings given were

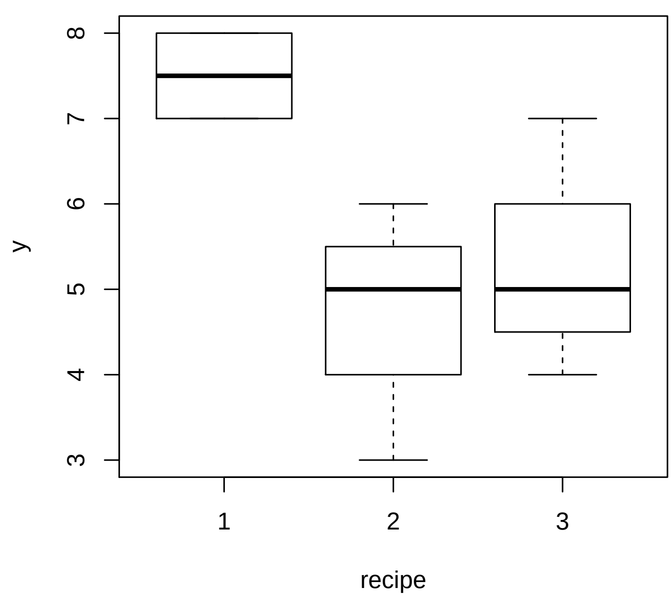
Recipe 1	Recipe 2	Recipe 3
7	3	5
8	6	7
7	5	4
8	5	5

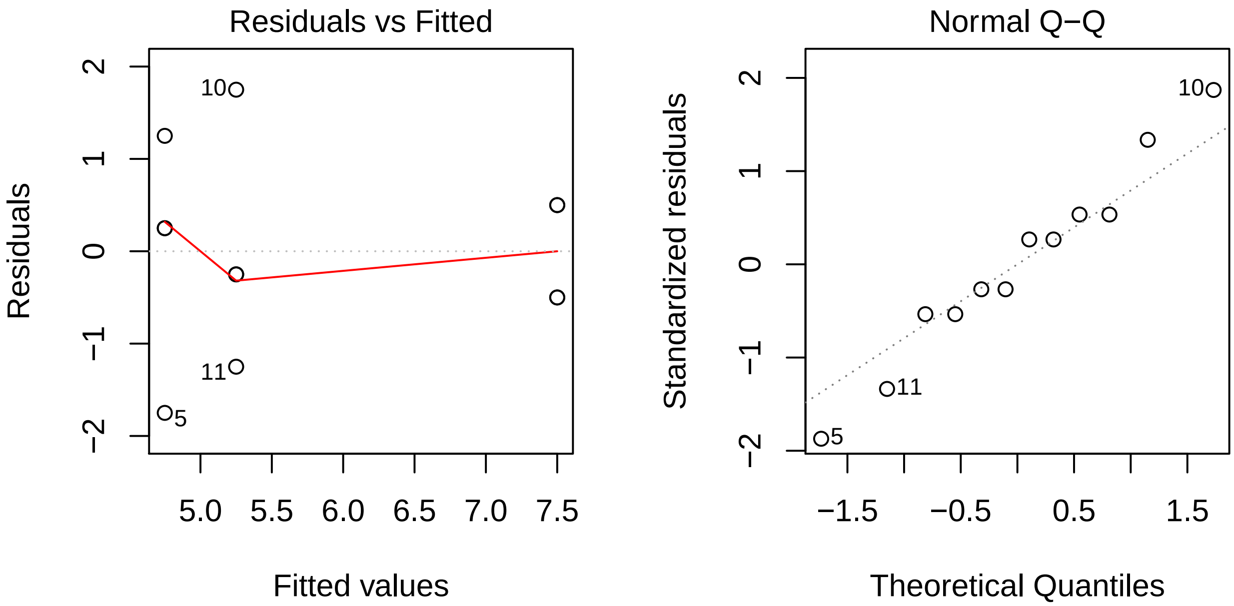
From these data, Mr. KG would like to know if it makes any difference which recipe he sends, and, if possible, which recipe would most please those cooking for the first time with a carbon steel wok.

Here is some R output:

```
    Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value  Pr(>F)
recipe     2 17.167  8.5833  7.3571 0.01278 *
Residuals  9 10.500  1.1667
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```





(a) Give the null and alternate hypotheses of interest to Mr. KG.

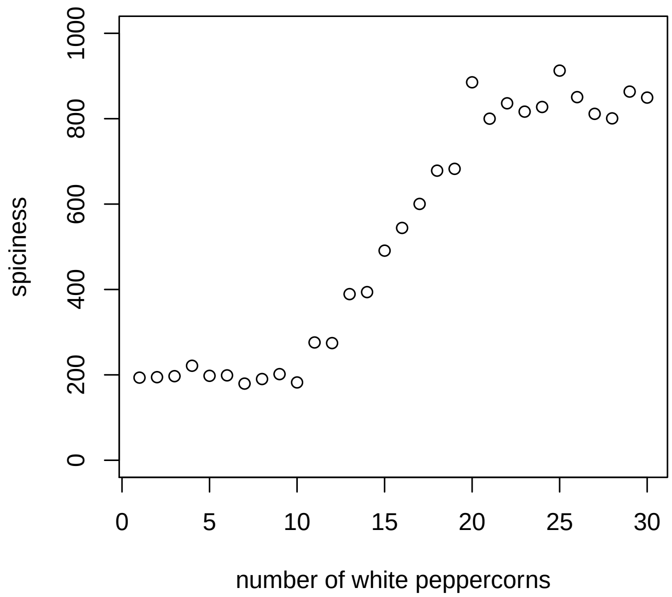
(b) What does the ratio given by 8.5833 / 1.1667, which appears in the ANOVA table, describe?

(c) What is the purpose of the *Residuals vs Fitted* plot?

(d) Comment on whether you think the ANOVA assumptions are satisfied.

(e) Write a data analysis report for Mr. KG explaining to him what he may conclude from the data. Give justifications for your claims.

6. A colleague of Mr. KG complains about the recipe selected on the basis of the study described in the previous question, saying that it is too spicy, owing to the amount of white pepper it calls for. In character with his perfectionism and penchant for statistical rigor, Mr. KG prepares the recipe 30 times, each time with a different number of peppercorns between 1 and 30. The spiciness of the dish is each time rated by his colleague (who does not know how many peppercorns were used) on a scale of 0 to 1000. The study resulted in the data plotted here:



A simple linear regression model is fit to the data. Below is some R output:

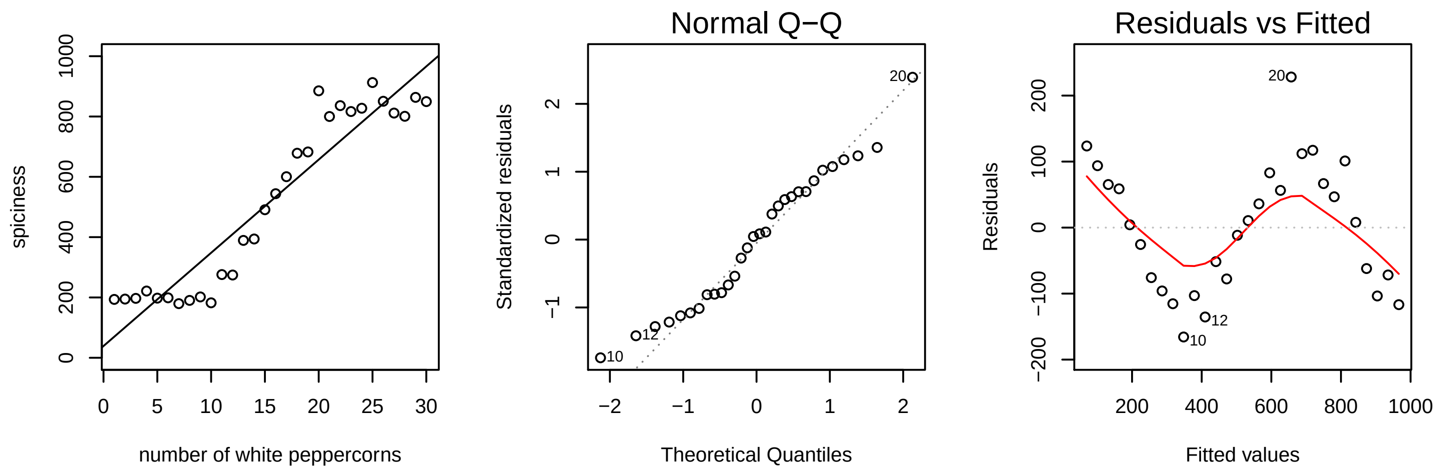
```
Call:
lm(formula = Y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-165.71  -77.27    6.13   66.35  228.19

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   38.983     36.503   1.068    0.295
x             30.906      2.056  15.031  6.2e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 97.48 on 28 degrees of freedom
Multiple R-squared:  0.8897,Adjusted R-squared:  0.8858
```

```
F-statistic: 225.9 on 1 and 28 DF,  p-value: 6.2e-15
```



(a) Give the intercept and slope of the least-squares regression line.

(b) State whether the assumptions of the linear regression model are satisfied for these data. Explain why they are or are not satisfied.

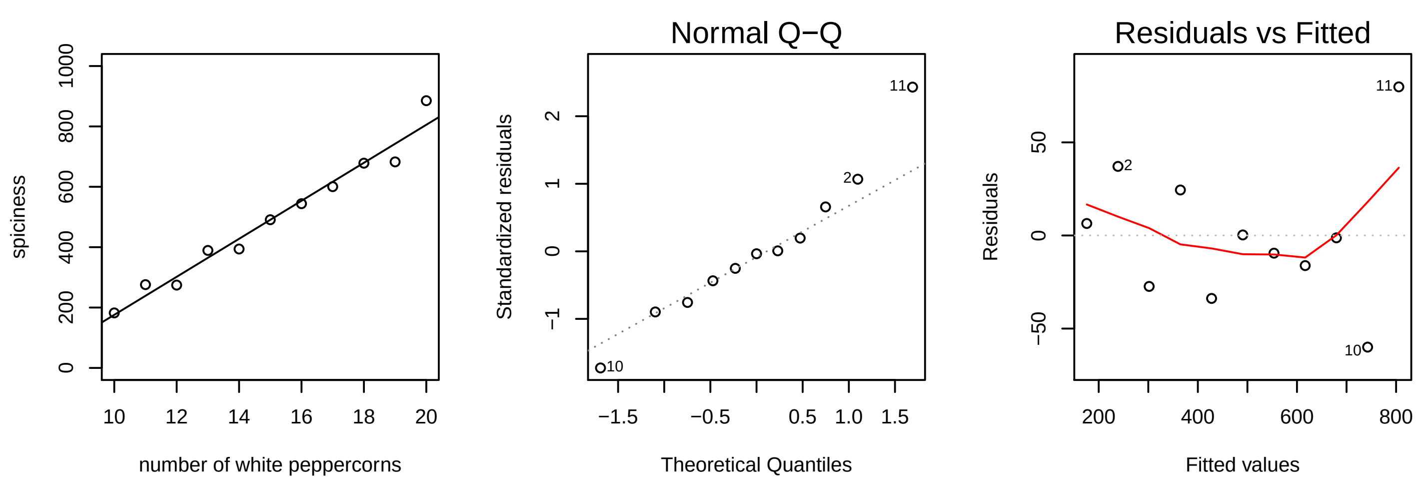
Mr. KG decides to focus on the relationship between the spiciness level and the number of peppercorns while the latter is between 10 and 20. Ignoring the part of the data with a number of peppercorns outside of this range, the above R output becomes:

```
Call:
lm(formula = Y[10:20] ~ x[10:20])

Residuals:
    Min      1Q  Median      3Q     Max
-59.956 -21.764  -1.273  15.438  79.812

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -453.711     58.109  -7.808 2.69e-05 ***
x[10:20]      62.960      3.791  16.609 4.64e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 39.76 on 9 degrees of freedom
Multiple R-squared:  0.9684,Adjusted R-squared:  0.9649
F-statistic: 275.9 on 1 and 9 DF,  p-value: 4.641e-08
```



(c) In the second analysis, what is the interpretation of the value 62.960 appearing in the R output?

(d) In the second analysis the value R^2 is reported as 0.9684. Give a careful interpretation of this value.

7. Mr. KG, interested in whether his customers tend to buy wok spatulas along with their woks, draws a sample of 100 recent orders from KG's Discount Store. Among the 100 orders sampled, 50 orders included a wok and 20 orders included a wok spatula. Of the 50 orders that included a wok, 15 also included a wok spatula.

(a) Fill out the counts in the table below, including row and column totals, to summarize the 100 sampled orders.

		Wok		
		Yes	No	
Wok spatula	Yes			
	No			

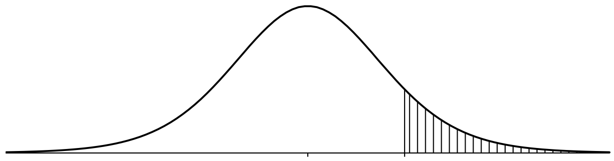
(b) Give the table of expected counts under the null hypothesis of no association.

(c) The output of the chi-squared test for association is given below. Based on the output, write for Mr. KG your conclusion concerning an association between buying a wok and buing a wok spatula.

```
Pearson's Chi-squared test

data:  M
X-squared = 6.25, df = 1, p-value = 0.01242
```

Upper tail probabilities of t-distributions



ν	α 0.100	0.050	0.025	0.010	0.005	0.001	0.0005
1	3.0777	6.3138	12.7062	31.8205	63.6567	318.3088	636.6192
2	1.8856	2.9200	4.3027	6.9646	9.9248	22.3271	31.5991
3	1.6377	2.3534	3.1824	4.5407	5.8409	10.2145	12.9240
4	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
21	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272	3.8193
22	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7921
23	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668	3.7454
25	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502	3.7251
26	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350	3.7066
27	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210	3.6896
28	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082	3.6739
29	1.3114	1.6991	2.0452	2.4620	2.7564	3.3962	3.6594
30	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852	3.6460
31	1.3095	1.6955	2.0395	2.4528	2.7440	3.3749	3.6335
32	1.3086	1.6939	2.0369	2.4487	2.7385	3.3653	3.6218
33	1.3077	1.6924	2.0345	2.4448	2.7333	3.3563	3.6109
34	1.3070	1.6909	2.0322	2.4411	2.7284	3.3479	3.6007
35	1.3062	1.6896	2.0301	2.4377	2.7238	3.3400	3.5911
40	1.3031	1.6839	2.0211	2.4233	2.7045	3.3069	3.5510
60	1.2958	1.6706	2.0003	2.3901	2.6603	3.2317	3.4602
120	1.2886	1.6577	1.9799	2.3578	2.6174	3.1595	3.3735
∞	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902	3.2905