STAT 481 **Midterm I-review** **Spring 2016**

1. **1.** Let $\overline{X}$ be the mean of a random sample of size n=36 from $N(\mu, \sigma^2=9)$. Our objective rule is to test H_0: $\mu=40$ against H_1: $\mu>40$. (1). Given the significance level $\alpha=0.05$, find its rejection region. (2). Find the probability of Type II error, β, and the power of the test at $\mu=40.5$? (3). Give a 95% confidence interval for the mean μ if $\overline{X}=40.8$.
2. **2.** A public opinion poll surveyed a simple random sample of 1000 voters. Respondents were classified by gender (male or female) and by voting preference (Republican, Democrat, or Independent). Results are shown in the contingency table below.

	Voting Preferences			Row total
	Republican	**Democrat**	**Independent**	
Male	200	150	50	400
Female	250	300	50	600
Column total	450	450	100	1000

Is there a gender gap? Do the men's voting preferences differ significantly from the women's preferences? Use a 0.05 level of significance.

3. **3.** The following are the burning time of chemical flares of two different formulations. It is assumed that both burning time follow normal distributions. The design engineers are interested in both mean and variances of the burning times.

	Observations										Sample Mean	Standard Deviation
Type I	65	82	81	67	57	59	66	75	82	70	70.4	9.264
Type II	64	56	71	69	83	74	59	82	65	79	70.2	9.367

(1). The engineers would like to test if the two variances are equal, state the appropriate null and alternative hypothesis.
(2). Use $\alpha=0.10$ and draw conclusions about the equality of the variances.
(3). If it is assumed that the burning times of the two different types have the same population variance (unknown σ^2), calculate an estimate of σ^2 by using the data from both samples.
(4). What's your test result about the comparison between two times, i.e. if there is significant difference between the means of the burning times of the two types?

4. An auto racing driver is given training and his timings on five different tracks are then measured. The times before and after training are given below.

Track	Before	After
1	226	198
2	706	701
3	559	589
4	975	892
5	280	264

Difference = Before – After

Sample mean of Difference: 20.4
Sample standard deviation of Difference: 41.15

Was the training beneficial to this driver, at $\alpha = 0.01$?

5. A shipping company offers customers the opportunity to purchase damage insurance for shipped packages. The shipping company is interested in whether or not a simple linear predictive model for package value can be constructed based on the package weight. Eight packages have been randomly sampled, and their weight (in pound) and value (in dollar) recorded in the following table:

Weight (X)	Value (Y)
51.67	34.22
106.08	78.48
86.28	45.86
93.74	51.03
94.46	46.06
98.96	60.34
74.76	41.06
81.38	47.44

$$\bar{x} = 85.92, \bar{y} = 50.56$$

$$\sum_{i=1}^{8}(x_i - \bar{x})^2 = 2{,}028.90$$

$$\sum_{i=1}^{8}(x_i - \bar{x})(y_i - \bar{y}) = 1333.78$$

$$\sum_{i=1}^{8}(y_i - \bar{y})^2 = 1{,}284.72$$

(1). Write down the simple linear regression model and model assumption.
(2). What is the objective function to obtain the least square estimators? Derive the normal equation for the linear coefficients in the model.
(3). Calculate the least squares estimates for the intercept β_0 and slope β_1.
(4). Compute the fitted value and the residual at the first and the last observations.
(5). Find SSR and SSE.

Brief Keys:

1. (1). Rejection region $\{\overline{X} > 40.82\}$; (2). $\beta(\mu = 40.5) = 0.74$, power$(\mu = 40.5) = 0.26$;
(3). 95% C.I. for μ: (39.82, 41.78).

2. H_0: Gender and voting preferences are independent. vs H_1: Gender and voting preferences are dependent. Statistics χ^2 = 16.2, p-value = 0.0003<0.05.

3.(1). H_0: $\sigma_1^2 = \sigma_2^2$ vs H_1: $\sigma_1^2 \neq \sigma_2^2$; (2). $F_o = 0.98$, p-value > 0.10;
(3). $S_p^2 = 9.3156^2$; (4). H_0: $\mu_1 = \mu_2$ vs H_1: $\mu_1 \neq \mu_2$, $t_o = 0.048$, p-value = 0.962.

4. No significant improvement, p-value = 0.165.

5. (1). $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, where i.i.d. errors $\varepsilon_i \sim N(0, \sigma^2)$.(3). $\hat{\beta}_0 = -5.92$, $\hat{\beta}_1 = 0.657$.
(4). $(28.027, 6.193), (47.547, -0.107)$.(5). $SSR = 875.77$, $SSE = 408.95$.