Do not write here!

Name:
Perm Number:

1

PSTAT 126 – B

Final exam

- Use a **DARK** pen or pencil, and write **INSIDE** the answer boxes provided.
- Write your name and perm number **CLEARLY** at the top of **EVERY** page, inside the boxes provided.
- No collaboration allowed, no electronic devices, no messanger pigeons.
- Please locate the nearest emergency exit. In some cases, your nearest exit may be behind you.

Part I: True or False

Please choose whether each of the following statement is TRUE or FALSE, and provide a short explanation or proof:

Question 1 (3pts) In a simple linear regression, $y = \beta_0 + \beta_1 x + \epsilon$, R^2 determines the sign of the slope of the regression line.

◯True
◯False

Answer:

Question 3 (3pts) The assumption that the random errors are normally distributed is not a necessary assumption to show that the ordinary least squares estimate is unbiased.

◯True
◯False

Answer:

2

PSTAT 126 – B

Name:
Perm Number:

Do not write here!

Part II: Multiple choice

Question 4 (1pt) Assume the following regression model, adjusted for 100 individuals:

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i,$$

which can be written in matrix for as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

What is the dimension of $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$?

◯ 100×3
◯ 100×2
◯ 3×3
◯ 2×2

Question 5 (1pt) When interpreting the coefficients in a multiple linear regression model, which of the following statements is true?

◯ A positive coefficient for an independent variable indicates that an increase in that variable is associated with an increase in the dependent variable

◯ A negative coefficient for an independent variable indicates that an increase in that variable is associated with a decrease in the dependent variable

◯ The magnitude of the coefficient indicates the strength of the relationship between the independent variable and the dependent variable

◯ All of the above

Question 6 (1pt) In a multiple linear regression model with k independent variables, which of the following is the formula for the adjusted R^2 value?

◯ $1 - (1 - R^2) \times (n - 1)/(n - k)$
◯ $R^2 \times (n - 1)/k$
◯ $(1 - R^2) \times (n - k)/(n - 1)$
◯ R^2/k

Do not
write here!

Name:

Perm Number:

3

PSTAT 126 – B

Question 8 (1pt) Which of the following statements is FALSE about the least squares estimate:

◯The LS estimate $\hat{\beta}$ is the most precise estimator among all the unbiased estimators of β.

◯The assumption that the random errors in the model are normally distributed is not needed to derive the least square estimates.

◯Least squares estimation can not be performed in the presence of categorical predictors.

4

PSTAT 126 − B

Name:
Perm Number:

Do not write here!

Part III: Proofs and derivations

Question 11 (7pts) As you remember, the sum of squared errors is $SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$. We know that in the Gausian-noise simple linear regression model, the ratio $SSE/\sigma^2 \sim \chi^2_{n-2}$.

(a) Given σ^2 and a number $\alpha > 0$, find a formula for an interval which contains SSE with probability $1-\alpha$, i.e., numbers l and u such that

$$\Pr(l \leq SSE \leq u) = 1 - \alpha$$

Express your answer in terms of n, σ^2, and the quantiles of χ^2 distributions.

Derivations:

Final answer:

(b) Using your answer from (a), find a formula for a $(1-\alpha)100\%$ confidence interval for σ^2. Your upper and lower limits should be expressed in terms of SSE, n, and the quantiles of χ^2 distributions.

Derivations:

Final answer:

Do not write here!

Name:

Perm Number:

5

PSTAT 126 – B

Question 12 (7pts) Consider the multiple linear regression model:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

(a) Prove that $E(\hat{\epsilon}_k) = 0$.

Derivations:

6

PSTAT 126 – B

Name:

Perm Number:

Do not write here!

Question 13 (6pts) Consider the simple linear regression model:

$$y = \beta_0 + \beta_1 x + \epsilon \quad \epsilon \sim N(0, \sigma^2)$$

Calculate $Var(\hat{\beta}_0 + \hat{\beta}_1)$, where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the LS estimates of the regression coefficients β_0 and β_1 respectively.

Derivations:

Final answer:

Do not write here!

Name:

Perm Number:

PSTAT 126 – B

	20	21	40	41	42	43	44	45	84	86	364	366
0.005	7.43	8.03	20.7	21.4	22.1	22.9	23.6	24.3	54.4	56.0	298	300
0.025	9.59	10.30	24.4	25.2	26.0	26.8	27.6	28.4	60.5	62.2	313	315
0.05	10.90	11.60	26.5	27.3	28.1	29.0	29.8	30.6	63.9	65.6	321	323
0.1	12.40	13.20	29.1	29.9	30.8	31.6	32.5	33.4	67.9	69.7	330	332
0.9	28.40	29.60	51.8	52.9	54.1	55.2	56.4	57.5	101.0	103.0	399	401
0.95	31.40	32.70	55.8	56.9	58.1	59.3	60.5	61.7	106.0	109.0	409	412
0.975	34.20	35.50	59.3	60.6	61.8	63.0	64.2	65.4	111.0	114.0	419	421
0.995	40.00	41.40	66.8	68.1	69.3	70.6	71.9	73.2	121.0	124.0	437	439

Table 1: Selected quantiles of χ^2 distributions: the probabilities are given by the rows, and the number of degrees of freedom by the columns

	20	21	40	41	42	43	44	45	84	86	364	366	Inf
0.005	-2.85	-2.83	-2.70	-2.70	-2.70	-2.70	-2.69	-2.69	-2.64	-2.63	-2.59	-2.59	-2.58
0.025	-2.09	-2.08	-2.02	-2.02	-2.02	-2.02	-2.02	-2.01	-1.99	-1.99	-1.97	-1.97	-1.96
0.05	-1.72	-1.72	-1.68	-1.68	-1.68	-1.68	-1.68	-1.68	-1.66	-1.66	-1.65	-1.65	-1.64
0.1	-1.33	-1.32	-1.30	-1.30	-1.30	-1.30	-1.30	-1.30	-1.29	-1.29	-1.28	-1.28	-1.28
0.9	1.33	1.32	1.30	1.30	1.30	1.30	1.30	1.30	1.29	1.29	1.28	1.28	1.28
0.95	1.72	1.72	1.68	1.68	1.68	1.68	1.68	1.68	1.66	1.66	1.65	1.65	1.64
0.975	2.09	2.08	2.02	2.02	2.02	2.02	2.02	2.01	1.99	1.99	1.97	1.97	1.96
0.995	2.85	2.83	2.70	2.70	2.70	2.70	2.69	2.69	2.64	2.63	2.59	2.59	2.58

Table 2: Selected quantiles of t distributions, with selected degrees of freedom; the last column gives quantiles of the Gaussian distribution.

F Table for Alpha 0.05

DF2	1	2	3	4	5	6	7	8	9	10	12
1	161.4476	199.5000	215.7073	224.5832	230.1619	233.9860	236.7684	238.8827	240.5433	241.8817	243.9060
2	18.5128	19.0000	19.1643	19.2468	19.2964	19.3295	19.3532	19.3710	19.3848	19.3959	19.4125
3	10.1280	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035
50	4.0343	3.1826	2.7900	2.5572	2.4004	2.2864	2.1992	2.1299	2.0734	2.0261	1.9515
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401	1.9926	1.9174
70	3.9778	3.1277	2.7355	2.5027	2.3456	2.2312	2.1435	2.0737	2.0166	1.9689	1.8932
80	3.9604	3.1108	2.7188	2.4859	2.3287	2.2142	2.1263	2.0564	1.9991	1.9512	1.8753
100	3.9361	3.0873	2.6955	2.4626	2.3053	2.1906	2.1025	2.0323	1.9748	1.9267	1.8503
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588	1.9105	1.8337
1000000	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522