MATH 668 EXAM 1 (March 3, 2016: 9:30am–10:45am) NAME ________________________

The exam is closed book; students are permitted to prepare one 8.5×11 page of formulas, notes, etc. that can be used during the exam. A calculator is permitted but not necessary for the exam.

Do 4 out of the 5 problems (20 points each, 80 points total). Clearly indicate the problem that you are omitting; if it is not clear, then the first 4 problems will be graded.

Problem 1. (20 points) Suppose that $\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is an n-dimensional column vector of outputs,

$\boldsymbol{X} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$ is an $n \times 2$ design matrix of fixed inputs, and $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ is a 2-dimensional column vector of coefficients.

(a - 6 pts) Let $Q(\boldsymbol{\beta}) = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2$. Compute $\dfrac{\partial Q}{\partial \boldsymbol{\beta}}$.

(b) Now suppose that $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$ where $\boldsymbol{e}$ follows a Normal distribution with mean vector $\boldsymbol{0}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and covariance matrix $\sigma^2 \boldsymbol{I}_2 = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ and that there is no real number c such that $x_i = c$ for all i.

(i - 4 pts) Let $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$ denote the maximum likelihood estimator of $\boldsymbol{\beta}$ and let $\boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$ be a 2-dimensional vector of fixed constants. What is the distribution of $\boldsymbol{a}^T\hat{\boldsymbol{\beta}}$?

(ii - 2 pts) Let $s^2 = \dfrac{\boldsymbol{r}^T\boldsymbol{r}}{n-2}$ denote the unbiased estimator of σ^2, where $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$. What is the distribution of $\dfrac{(n-2)s^2}{\sigma^2}$?

(iii - 4 pts) Prove that $\hat{\boldsymbol{\beta}}$ and $\boldsymbol{r}^T\boldsymbol{r}$ are independent.

(iv - 4 pts) Prove that $\dfrac{\boldsymbol{a}^T\hat{\boldsymbol{\beta}} - \boldsymbol{a}^T\boldsymbol{\beta}}{s\sqrt{\boldsymbol{a}^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{a}}}$ follows a $t(n-2)$ distribution.

Problem 2. (20 points) Suppose that $\sum_{i=1}^{10} x_i = 0$, $\sum_{i=1}^{10} y_i = 0$, $\sum_{i=1}^{10} x_i^2 = 6$, $\sum_{i=1}^{10} y_i^2 = 16$, and $\sum_{i=1}^{10} x_i y_i = 6$. Assume the simple linear regression model where y_i are independent Normal$(\beta_0 + \beta_1 x_i, \sigma^2)$ random variables for $i = 1, \ldots, 10$. Perform a test of $H_0 : \beta_1 = 2$ versus $H_A : \beta_1 < 2$ at level .05. Make sure to state your conclusion and to show sufficient details (such as your test statistic and the corresponding critical value or bound for the P-value) to justify your conclusion.

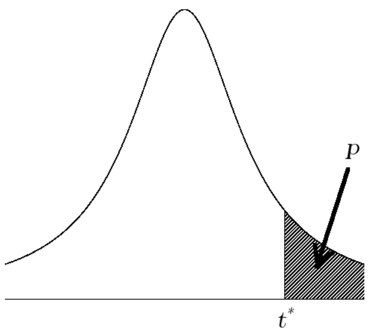
Problem 3. (20 points)
(a - 10 pts) The moment generating function of a random variable with a $\chi^2(\nu)$ distribution is $M(t) = \dfrac{1}{(1-2t)^{\nu/2}}$.
Suppose that $q_1, \ldots, q_k$ are independent χ^2 random variables with $\nu_1, \ldots, \nu_k$ degrees of freedom, respectively.
Prove that $q = \sum_{i=1}^{k} q_i$ has a $\chi^2\left(\sum_{i=1}^{k} \nu_i\right)$ distribution.
(b - 10 pts) Suppose that $\boldsymbol{y}$ is an m-dimensional multivariate Normal$(\boldsymbol{0}, \boldsymbol{V})$ random vector. Also, let $\lambda_1, \ldots, \lambda_m$ be eigenvalues of the covariance matrix $\boldsymbol{V}$. Let $\boldsymbol{V} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^T$ where $\boldsymbol{\Lambda}$ is a diagonal matrix with diagonal elements $\lambda_1, \ldots, \lambda_m$ and $\boldsymbol{T}$ is an $m \times m$ matrix such that $\boldsymbol{T}^T\boldsymbol{T} = \boldsymbol{I}$.
If $\boldsymbol{V}$ is invertible, then prove that $\boldsymbol{y}^T\boldsymbol{V}^{-1}\boldsymbol{y}$ has a $\chi^2(m)$ distribution.

Problem 4. (20 points) Suppose that $\boldsymbol{y}$ is a 5-dimensional Normal$(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I})$ random vector where $\boldsymbol{X}$ is a 5×2 matrix with columns $\boldsymbol{J}$ and $\boldsymbol{x}$, $\boldsymbol{J}$ is an 5-dimensional vector of ones, $\boldsymbol{x}$ is a non-random 5-dimensional vector, $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ is a unknown vector of intercept and slope parameters, and σ^2 is the unknown variance parameter.
Let $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\sigma}^2$ denote the maximum likelihood estimators of β_0, β_1 and σ^2, respectively.
If $\boldsymbol{J}^T\boldsymbol{x} = 3$ and $\boldsymbol{x}^T\boldsymbol{x} = 2$, find values A, B, and C such that

$$A\left(\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}}\right)^2 + B\left(\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}}\right)\left(\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}}\right) + C\left(\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}}\right)^2 \leq 1$$

is a 95% confidence ellipse for $\boldsymbol{\beta}$.

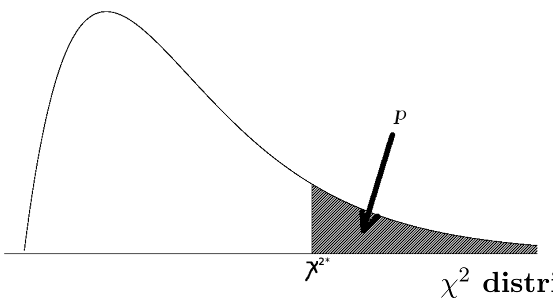
Problem 5. (20 points) Let $\mathcal{X}_k = \begin{bmatrix} 1 \\ x_k \end{bmatrix}$ for $k = 1, \ldots, n$, let $\boldsymbol{X}$ be an $n \times 2$ matrix with kth row $\mathcal{X}_k^T$, and let $\boldsymbol{y}$ be an n-dimensional vector where y_k denotes its kth element.
(a - 8 pts) Show that $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\mathcal{X}_n = \begin{bmatrix} \dfrac{1}{n} - \dfrac{\bar{x}(x_n - \bar{x})}{S_{xx}} \\ \\ \dfrac{x_n - \bar{x}}{S_{xx}} \end{bmatrix}$ where $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ and $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$.
(b - 4 pts) Compute $h_{nn} = \mathcal{X}_n^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\mathcal{X}_n$, the nth diagonal element of the hat matrix.
(c - 8 pts) Let $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$ be the minimizer of $Q(\beta_0, \beta_1) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$ and let $\hat{\boldsymbol{\beta}}_{(n)} = \begin{bmatrix} \hat{\beta}_{0,(n)} \\ \hat{\beta}_{1,(n)} \end{bmatrix}$ be the minimizer of $Q(\beta_0, \beta_1) = \sum_{i=1}^{n-1}(y_i - \beta_0 - \beta_1 x_i)^2$. Show that $\hat{\beta}_{1,(n)} = \hat{\beta}_1 - \dfrac{(y_n - \hat{\beta}_0 - \hat{\beta}_1 x_n)(x_n - \bar{x})}{\left(1 - \frac{1}{n}\right)S_{xx} - (x_n - \bar{x})^2}$.



The critical value t^* is the value such that the area under the density curve of a t distribution with *df* degrees of freedom to the right of t^* is equal to p. It is also the value such that the area under the curve between $-t^*$ and t^* is equal to C.

t distribution critical values

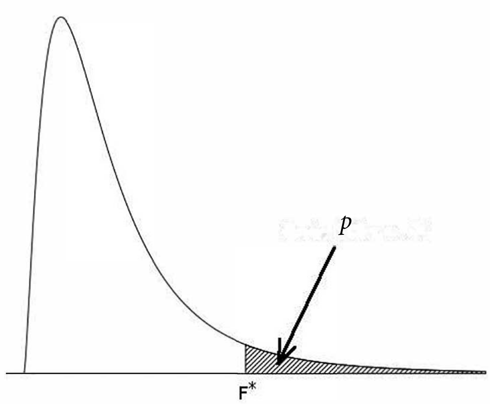
	Upper-tail probability p				
df	.10	.05	.025	.01	.005
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
	80%	90%	95%	98%	99%
	Confidence level C				



The critical value χ^{2*} is the value such that the area under the density curve of a χ^2 distribution with df degrees of freedom to the right of χ^{2*} is equal to p.

χ^2 distribution critical values

	Upper-tail probability p				
df	.10	.05	.025	.01	.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75



The critical value F^* is the value such that the area under the density curve of an F distribution with $df1$ degrees of freedom in the numerator and $df2$ degrees of freedom in the denominator to the right of F^* is equal to p.

F distribution critical values

		$p = .05$				
		$df1$				
		1	2	3	4	5
	1	161.45	199.5	215.71	224.58	230.16
	2	18.51	19.00	19.16	19.25	19.30
$df2$	3	10.13	9.55	9.28	9.12	9.01
	4	7.71	6.94	6.59	6.39	6.26
	5	6.61	5.79	5.41	5.19	5.05

		$p = .025$				
		$df1$				
		1	2	3	4	5
	1	647.79	799.50	864.16	899.58	921.85
	2	38.51	39.00	39.17	39.25	39.30
$df2$	3	17.44	16.04	15.44	15.10	14.88
	4	12.22	10.65	9.98	9.60	9.36
	5	10.01	8.43	7.76	7.39	7.15