Math 326 Section 01
Professor Carty
60 points possible

Exam 2
April 15, 2022

Name: Matthew Wilder

1. No hats or dark sunglasses. All hats are to be removed.
2. All book bags are to be closed and placed in a way that makes them inaccessible. Do not reach into your bag for anything during the exam. If you need extra pencils, pull them out now.
3. You are allowed a one-page handwritten sheet of formulas. Notes, theorems, lemmas, etc. are not allowed on the formula sheet. Please put your name on your formula sheet and hand it in with the exam.
4. Be sure to print your proper name clearly.
5. Calculators are optional, but if you choose to use one you may ONLY use the battery or solarpowered Texas Instruments BA35 model calculator, the BA II Plus, the BA II Plus Professional, the TI30Xa or TI30X II (IIS solar or IIB battery), or TI-30X MultiView (XS Solar or XB Battery).
6. All electronic devices, including cell phones and other wearable devices, must be silenced or powered off and stored out of sight for the entirety of the exam.
7. If you have a question, raise your hand and I will come to you. Once you stand up, you are done with the exam. If you have to use the facilities, do so now. You will not be permitted to leave the room and return during the exam.
8. Every exam is worth a total of **60 points**. Including the cover sheet, each exam has 6 pages.
9. At 11:45, you will be instructed to put down your writing utensil. Anyone seen continuing to write after this announcement will have their exam marked and lose all points on the page they are writing on. At this time, you will use your phone to take PDFs of the pages and upload them into Canvas as a single PDF document (the same way you do with homework.) Solutions must be uploaded by 10:55.
10. If you finish early, quietly and respectfully perform the preceding tasks. You may leave early.
11. You will hand in the paper copy of the exam on your way out of the classroom.
12. You have forty-five minutes to complete the exam. I hope you do well.

1. (10 points) The manufactures of a cold medicine want to test a new version of their product to see if the new version is better. Volunteer cold sufferers were randomly assigned to the old version (Drug #1) or to the new version (Drug # 2). Let p_1 denote the percentage of all cold sufferers whose symptoms improved using Drug #1 and p_2 be the percentage who improved using Drug #2. Among the test groups, 130 of 200 using Drug #1 improved, while 252 of 300 using Drug #2 were improved.

(a) What are the hypotheses to be tested? What is the test statistic?

$p_1 = 0.65$
$p_2 = 0.84$
$n_1 = 200$
$n_2 = 300$

$H_0 : p_1 = p_2$

$H_a : p_1 < p_2$ (p_2 better)

$$T = \frac{\bar{X}_1 - \bar{X}_2 - 0}{S_p\sqrt{\frac{1}{200} + \frac{1}{300}}}$$

$$S_p = \sqrt{\frac{(199)S_1^2 + (299)S_2^2}{498}}$$

(b) Calculate the test statistic and the associate P-value.

(c) At the $\alpha = 0.05$ level, what is your conclusion? What is the practical conclusion to the drug manufacturer?

$Z_{0.025} = 1.960$

Since $p > \alpha$ we Can conclude that drug 2 is better, aka reject the null.

2. (10 points) Suppose that we want to test the hypotheses $H_0 : \mu = 7$ versus $H_a : \mu > 7$ in a normal distribution based on a random sample of size $n = 4$ with rejection region $\bar{Y} > 9$. Suppose that the standard deviation σ is known to be 3.0.

(a) Find α, the probability of Type I error.

P(reject H_0 when H_0 is true)

(b) Find β, the probability of Type II error, if $\mu = 10$.

3. (15 points) Suppose that $X_1, \ldots, X_n$ is a iid sample from an exponential distribution with parameter $\theta > 0$ unknown:

$$f(y) = e^{-x/\theta}/\theta, 0 < x < \infty.$$

(a) Find and simplify the likelihood function $L(x_1, x_2, \ldots, x_n|\theta)$ and determine a sufficient statistic for θ.

$$L = \prod_{i=1}^{n} e^{-x/\theta}/\theta = \frac{1}{\theta^n} e^{\Sigma -\frac{x}{\theta}} = \boxed{\frac{1}{\theta^n} e^{-\frac{1}{\theta}\Sigma x}}$$

$$\boxed{S = \Sigma x}$$

(b) Determine the rejection region for the most powerful test of $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$, assuming $\theta_a > \theta_0$. (You do not have to determine how your statistic is distributed.)

$$\frac{L(\theta_0)}{L(\theta_a)} \leq k \text{ in } \overset{\text{critical}}{\text{region}}. \qquad \frac{\frac{1}{\theta_0^n} \exp\left[-\frac{1}{\theta_0}\Sigma x\right]}{\frac{1}{\theta_a^n} \exp\left[-\frac{1}{\theta_a}\Sigma x\right]} > k$$

$$\frac{\theta_a^n}{\theta_0^n} \exp\left[-\frac{1}{\theta_0}\Sigma x + \frac{1}{\theta_a}\Sigma x\right] = \left(\frac{\theta_a}{\theta_0}\right)^n \exp\left[\Sigma x\left(\frac{1}{\theta_a} - \frac{1}{\theta_0}\right)\right] > k$$

is the RR

(c) Is the test you defined in part (a) uniformly most powerful for the alternative $\theta > \theta_0$? Briefly explain your answer.

Yes it does not depend on a choice of θ

4. (25 points) Suppose that we have a designed experiment to investigate the effects of 2 factors (x_1 and x_2) on an outcome y. We consider the model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon.$$

The data is in the following table:

x_1	-1	-1	-1	0	0	0	1	1	1
x_2	-1	0	1	-1	0	1	-1	0	1
y	10	8	3	8	6	2	5	2	1

(a) Determine the matrices $\mathbf{X}$, $\mathbf{X}^T$, $\mathbf{X}^T\mathbf{X}$, $\mathbf{X}^T\mathbf{Y}$, and $(\mathbf{X}^T\mathbf{X})^{-1}$.

$$X = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} 10 \\ 8 \\ 3 \\ 8 \\ 6 \\ 2 \\ 5 \\ 2 \\ 1 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} \qquad X^T Y = \begin{bmatrix} 45 \\ -13 \\ -17 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/6 & 0 \\ 0 & 0 & 1/6 \end{bmatrix}$$

(b) Find the model to the data.

$$\hat{\beta} = (X^T X)^{-1}(X^T Y)$$

$$= \begin{bmatrix} 1 & & \\ & 1/6 & \\ & & 1/6 \end{bmatrix} \begin{bmatrix} 45 \\ -13 \\ -17 \end{bmatrix} = \begin{bmatrix} 45 \\ -13/6 \\ -17/6 \end{bmatrix} \equiv \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$\hat{y} = 45 - \frac{13}{6}x_1 - \frac{17}{6}x_2$$

(c) Determine the residual for the data point $(1, -1, \cancel{10}\ 5)$.

$$\hat{y}(1,-1) = 45 - \frac{13}{6} + \frac{17}{6} = 45 + \frac{4}{6} = 45 + \frac{2}{3}$$

actual: 5

$$\left|45 + \frac{2}{3} - 5\right| = 40 + \frac{2}{3} = \boxed{40.6\overline{7}}$$

(d) Use matrices to calculate SSE and determine S. (Note $\sum y^2 = 307$.)

oh...

$$SSE = Y^T Y - \hat{\beta}^T X^T Y$$

$\|Y\|^2 = 307$ (mth 307?)

$$\begin{bmatrix} 45 & \frac{-13}{6} & \frac{-17}{6} \end{bmatrix} \begin{bmatrix} 45 \\ -13/6 \\ -17/6 \end{bmatrix} = 45^2 + \left(\frac{13}{6}\right)^2 + \left(\frac{17}{6}\right)^2$$

$2025 + 169 + 289 = \frac{458}{36}$

$12.7\overline{2}$

$= 2037.722$

$$S = \sqrt{\frac{SSE}{n-(k+1)}} = \sqrt{\frac{SSE}{9-3}} \approx \sqrt{339.6}$$

≈ 18.429

$SSE = 2037.7\overline{2}$

$S \approx 18.429$

(e) Test $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 < 0$. Perform the test at the $\alpha = 0.01$ level. What can you conclude about the usefulness of the x_1 term in the model?

$$a^T\hat{\beta} \pm t_{\alpha/2}(6)\, s\sqrt{a^T(X^TX)^{-1}a}$$

$t_{0.005}$ 3.707

df 6

$$a = \begin{pmatrix} 45 \\ 0 \\ -17/6 \end{pmatrix}$$

$$T = \frac{a^T\hat{\beta} - (a^T\beta)_0}{s\sqrt{a^T(X^TX)^{-1}a}}$$

$t \cdot s$

SVD probably says its not necessary.
Smallest singular value (guess).

$$\begin{pmatrix} 45 \\ 0 \\ -17/6 \end{pmatrix}^T \begin{pmatrix} 45 \\ -13/6 \\ -17/6 \end{pmatrix} \quad 2033.028 = a^T\hat{\beta}$$

$2033.028 \pm$

(f) Find a 95% confidence interval for $E(Y)$ when $x_1 = 0.75$ and $x_2 = -1$.

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2}(6)\, S\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

$$t_{0.025}(6) = 2.447$$

Table entry for p and C is th...

Matthew Wilder

$\hat{p} \pm t_{\alpha/2}(df)\sqrt{\frac{\hat{p}\hat{q}}{n}}$

$\hat{y} = \hat{b}_0 + \hat{b}_1 x \quad \hat{b}_1 = \rho \frac{S_y}{S_x} \quad \hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}$

$\int_c^d f(g(x))\, g'(x)\, dx = \int_{g(c)}^{g(d)} f(x)\, dx$

$\text{power}(\theta) = 1 - \beta(\theta)$

$b = \frac{\langle y, 1\rangle}{n} - \left\langle m, \frac{\langle x, 1\rangle}{n}\right\rangle$

$\int u\,dv = uv - \int v\,du$

$m = \frac{Cov(x,y)}{S_x^2} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2}$

$(fg)' = fg' + f'g \quad \left(\frac{1}{g}\right)' = \frac{-g'}{g^2}$

$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

$\rho^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}}$

$(f/g)' = \frac{f'g - fg'}{g^2} \quad N(\mu, \sigma^2)$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - n\bar{x}\bar{y}$

$\sigma_1 = S_1 \quad \sigma_2 = S_2$

$-Z_* = \frac{C^* - \mu}{\sigma/\sqrt{n}} \iff C^* = \mu - Z_* \frac{\sigma}{\sqrt{n}}$

$m = \frac{\Sigma x_i y_i - n\bar{x}\bar{y}}{\Sigma x_i^2 - n\bar{x}^2} = \frac{n\langle x, y\rangle - \langle x, 1\rangle\langle y, 1\rangle}{n\langle x, x\rangle - \langle x, 1\rangle^2}$

$\frac{d}{dx}(\ln(x)) = \frac{1}{x}$

$\mu_0 \pm Z_* \frac{\sigma}{\sqrt{n}}$

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - n\bar{x}\bar{y}$

$\frac{d}{dx}(e^{ax}) = ae^{ax}$

$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - n\bar{x}^2$

$n = \frac{(Z_a + Z_b)^2 \sigma^2}{(\mu_a - \mu_0)^2}$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$E[Y] = \beta_0 + \beta_1 x$

$Y = \beta_0 + \beta_1 x + \varepsilon$

$E[\varepsilon] = 0 \quad V[Y] = V[\varepsilon] = \sigma^2$

$\frac{d}{dx}(a^x) = a^x \ln a$

$Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$

$SSE = \Sigma(y_i - \hat{y}_i)^2$

$\frac{d}{dx} \log_a x = \frac{1}{x \ln a}$

$\bar{x} \pm t_* \sqrt{s^2/n}$

$Var(\hat{\beta}_0) = \frac{\sigma^2 \Sigma x_i^2}{n S_{xx}}$

$s^2 = \frac{SSE}{n-2}$

$\ln(xy) = \ln x + \ln y$

$\ln(x/y) = \ln x - \ln y$

$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$

$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{x}\sigma^2}{S_{xx}}$

$SSE = S_{yy} - \hat{\beta}_1 S_{xy}$

$\ln x^a = a \ln x$

$S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - n\bar{y}^2$

$\dim V = \dim \text{range}\, T + \dim \text{null}\, T$

$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}(df)\left[S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$

$T = \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}}$

$e_j = \frac{v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}}{\|\ \ \text{top}\ \ \|}$

$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \quad df = n_1 + n_2 - 2$

$\hat{\beta}_i \pm Z_{\alpha/2}(df)\sqrt{Var(\hat{\beta}_i)}$

$n-2$

$|\langle u, v\rangle| \le \|u\|\|v\|$

$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$Var(\hat{\beta}_0) = \frac{\sigma^2 \Sigma x_i^2}{n S_{xx}} \sim \frac{s^2 \Sigma x_i^2}{n S_{xx}}$

$\arccos \frac{\langle x, y\rangle}{\|x\|\|y\|}$

$\|u + v\|^2 = \|u\|^2 + \|v\|^2$

$\langle x, y\rangle = \frac{\|x + y\|^2 - \|x - y\|^2}{4}$

$\|v\| = \sqrt{\langle v, v\rangle}$

$\alpha = Pr(C \mid \theta_0) = \int_C \Pi f(x_i \mid \theta_0)$

$C \subseteq x^n \subseteq R^n$

$(\hat{\beta}_0 + \hat{\beta}_1 x^*) \pm t_{\alpha/2}(n-2)\, s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$

$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$

$(\hat{\beta}_0 + \hat{\beta}_1 x^*) \pm t_{\alpha/2}(n-2)\, S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$

$\langle u, v\rangle = \overline{\langle u, v\rangle}$

- $P((x_1, \ldots, x_n) \in C \mid \theta_0) = \alpha$

$T = \frac{y_0 - (\hat{\beta}_0 + \hat{\beta}_1 x^*)}{S\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}}$

$(X^T X)^{-1} = \frac{1}{n S_{xx}} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix}$

$Var(X) = E(X^2) - E(X)^2$

$Var(\alpha X) = \alpha^2 Var(X)$

$Var(Y^*) = s^2\left(1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}\right)$

$= \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$

$E[cX] = cE[X]$

$E[c] = c$

$\vec{Y} = \beta_0 \vec{x}^0 + \vec{\beta_1 x} + \vec{\beta_2 x^2} + \cdots \beta_k \vec{x^k}$

$Z = \frac{\hat{B}_i - (B_i)_0}{\sqrt{Var(B_i)}}$

$XB = Y \quad X^T X \hat{\beta} = X^T Y$

$\sigma^2 (X^T X)^{-1} = \begin{bmatrix} Cov(B_0, B_0) & \cdots & Cov(B_0, B_n) \\ \vdots & \ddots & \\ Cov(B_n, B_0) & \cdots & Cov(B_n, B_n) \end{bmatrix}$

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$T = \frac{\hat{B}_i - (B_i)_0}{S\sqrt{c_{ii}}}$

$\hat{\beta} = (X^T X)^{-1} X^T y$

$SSE = Y^T Y - \hat{\beta}^T X^T Y$

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

$X^T Y = \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{bmatrix}$

$y = mx + b \iff x = \frac{y - b}{m}$

$[\ \ ][\ \ ] = \begin{bmatrix} \langle \text{row}_1, \text{col}_1\rangle & \cdots & \langle \text{row}_1, \text{col}_n\rangle \\ \vdots & & \vdots \\ \langle \text{row}_n, \text{col}_1\rangle & \cdots & \langle \text{row}_n, \text{col}_n\rangle \end{bmatrix}$

$Z_{0.05} = 1.645$

$Z_{0.025} = 1.960$

$Z_{0.005} = 2.576$

$Var(\hat{B}_i) = c_{ii}\sigma^2$

$Cov(\hat{B}_i, \hat{B}_j) = c_{ij}\sigma^2 \quad (X^T X)^{-1} \equiv c_{ij} \quad E(\hat{B}_i) = B_i$

$s^2 = \frac{SSE}{n - (k+1)} \quad \frac{(n - (k+1))s^2}{\sigma^2} \sim \chi^2(n - (k+1))$

$\sigma = \sqrt{\sigma^2}$

$Var(a^T \hat{\beta}) = a^T (X^T X)^{-1} a \cdot \sigma^2$

$S = \sqrt{\frac{SSE}{n - k - 1}}$

$(\quad)(\ ) = (\ ) \quad (\ )(\quad) = (\qquad)$

$T = \frac{a^T \hat{\beta} - (a^T B)_0}{S\sqrt{a^T (X^T X)^{-1} a}}$

$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$

$M = U \Sigma V^T$

$a^T \hat{\beta} \pm t_{\alpha/2}(n - k - 1)\, S\sqrt{a^T (X^T X)^{-1} a}$

$L(x_1, \ldots, x_n \mid \theta) = f(x_1 \mid \theta) \cdots f(x_n \mid \theta)$

det : + − + − + ...

$1 + 1 = 2$

$\int_a^b f = F(b) - F(a) \quad \frac{d}{dx}\left[\int_a^x f\right] = f(x)$

1 3 6 10 15 21 28 36