AMS 315 S2022 Midterm 2 Summary Statistics

Upper Quartile Point	315	Mean	255.1
Median	280	Standard Deviation	77.8
Lower Quartile Point	215	IQR SD	134

AMS 315 S2022 Midterm 2 Examination Grading Guidance

General guidance

1. My answers have a high number of digits. My logic is that the extra accuracy helps you to study. Answers that are correctly rounded off from my answer were accepted.
2. The points that you earned on each question were put at the top of each response. I identified the first error that I penalized.
3. I took off 2 points for minor computational errors and 5 points for more serious errors. An answer that had a substantive error in a calculation had a 20-point deduction. For example, forgetting to divide by the square root of *n* in a confidence interval.
4. A decision to accept or reject a null hypothesis that is inconsistent with the calculations of the problem had up to a 35-point deduction.

Guidance on specific problems

1. Confidence interval for ρ: +15 for correct F(r). An additional +10 for correct 99% CI for $F(\rho)$. -25 for incorrect transformations of the endpoints; -10 wrong $|z_\alpha|$; -50 for a wrong procedure.
2. Regression Problem: A: -15 for incorrect dependent variable—only deduct once for this error; that is no further deductions for consistent work; -15 each incorrect degree of freedom; -10 for each incorrect sum of squares—do not deduct for consistent errors; -20 inconsistent decision about hypothesis; -5 no table, but each entry of table given; B: -10 incorrect slope; -15 incorrect standard error of slope; -10 wrong $|z_\alpha|$; wrong $\hat{Y}(x)$ by combining estimated equations using IV and DV; C: -20 for confidence interval reported in prediction interval problem, and vice versa; -5 wrong argument in fitting function.
3. Partial correlations: -10 for each incorrect partial correlation; -10 for forgetting square root in denominator; -20 for correct partial correlations that are mislabeled.
4. Explanation or Mediation: 30 points for a correct answer; 0 for an incorrect answer. -30 for straddling: for example, saying in text explanation model and showing a mediation model in a diagram.
5. Two independent variable multiple regression anova: Standard anova table penalties: -15 for each incorrect degrees of freedom; -10 for each incorrect entry. Only penalize an error when it is made; do not penalize subsequent correct uses of an incorrect entry. Wrong variable used in Anova table -15; -25 do not divide by DFE in MSE calculation. No

decision or inconsistent decision about the hypothesis test: -35. Reversed sequence from what was asked: -25; -15 for $F(x, w)$ rather than $F(x|w)$.

6. Sample size in correlation study: +15 for correct $F(\rho_1)$; -15 for wrong $|z_\alpha|$; -15 for wrong $|z_\beta|$; -30 forget to square.
7. $vcv(MY)$: Give +15 points for correct M; -10 omitting σ^2; -20 for each incorrect variance; -10 for each incorrect covariance (the two covariances should be equal).

X2S2022P1AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1. A research team conducted a longitudinal study of participants between 40 and 50 years of age. They measured each participant's daily caloric consumption at age 40. They also measured systolic blood pressure at age 50. The team collected data on n = 582 participants. The Pearson product moment correlation coefficient between the two variables was 0.58. What is the 95% confidence interval for the correlation of daily caloric consumption at age 40 and systolic blood pressure at age 50? This problem is worth 50 points.

$$F(0.58) = \frac{1}{2} \ln \left(\frac{1+.58}{1-.58}\right) = \frac{1}{2} \ln \left(\frac{1.58}{0.42}\right) = \frac{1}{2} \ln (3.762)$$

$$= \frac{1}{2}(1.325) = 0.6625$$

$$\sigma(F(R)) \cong \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{579}} = 0.0416$$

95% CI FOR $F(\rho)$ IS: $0.6625 \pm 1.960(0.0416)$

$$= 0.6625 \pm 0.0815 = 0.581 \text{ TO } 0.7440$$

RIGHT END POINT OF 95% CI FOR ρ:

$$e^{2(0.7440)} = e^{1.488} = 4.428.$$

$$R_R = \frac{4.428-1}{4.428+1} = \frac{3.428}{5.428} = 0.632.$$

LEFT END POINT

$$e^{2(.581)} = e^{1.162} = 3.196.$$

$$R_L = \frac{3.196-1}{3.196+1} = \frac{2.196}{4.196} = 0.523$$

THE 95% CI FOR ρ FROM $R = 0.58$, $n = 582$

IS 0.523 TO 0.632

X2S2022P1BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

1 A research team conducted a longitudinal study of participants between 50 and 60 years of age. They measured each participant's daily caloric consumption at age 50. They also measured systolic blood pressure at age 60. The team collected data on n = 482 participants. The Pearson product moment correlation coefficient between the two variables was 0.43. What is the 99% confidence interval for the correlation of daily caloric consumption at age 50 and systolic blood pressure at age 60? This problem is worth 50 points.

$$F(0.43) = \frac{1}{2} \ln\left(\frac{1+.43}{1-.43}\right) = \frac{1}{2} \ln\left(\frac{1.43}{0.57}\right) = \frac{1}{2} \ln(2.509)$$

$$= \frac{1}{2}(0.9198) = 0.4599.$$

99% CI FOR $F(\rho)$ IS.

$$0.4599 \pm 2.576\sqrt{\frac{1}{482-3}}$$

$$0.4599 \pm 2.576(0.0457)$$

$$= 0.4599 \pm 0.1177 = 0.3422 \text{ TO } 0.5776$$

RIGHT END POINT OF 99% CI FOR ρ:

$$e^{2(.5776)} = e^{1.1552} = 3.175$$

$$r_R = \frac{3.175-1}{3.175+1} = \frac{2.175}{4.175} = 0.521.$$

LEFT END POINT:

$$e^{2(0.3422)} = e^{0.6844} = 1.983.$$

$$r_L = \frac{1.983-1}{1.983+1} = \frac{0.983}{2.983} = 0.3295.$$

THE 99% CI FOR ρ FROM $r = 0.43$, $n = 482$

IS 0.3295 TO 0.521.

2 A research team conducted a chemical response study of the activity of a biological cell associated with the amount of chemical input to the cell. The team collected data on $n = 125$ cells. The average chemical input was 1,021 units, with an observed standard deviation of 64.8 units (the divisor in the underlying variance calculation was $n - 1$). The average chemical response was 1,667 units, with an observed standard deviation of 445.8 units (the divisor in the underlying variance calculation was $n-1$). The Pearson product moment correlation coefficient between the two variables was 0.91. The research team sought to estimate the regression of the chemical response on the chemical input to the cell.

a. Complete the analysis of variance table for the regression of the chemical response on the chemical input to the cell. Test the null hypothesis that the slope of this regression is zero at levels of significance 0.10, 0.05, and 0.01. This part is worth 25 points.

b. Find the estimated regression of the regression of the chemical response on the chemical input to the cell. Find the 99% confidence interval for the slope. This part is worth 25 points.

c. Use the least-squares equation to estimate the chemical response to a chemical input of 1,200 units. What is the 99% prediction interval for the chemical response of a cell subjected to a chemical input of 1,200 units. This part is worth 20 points.

A. IV = INPUT DV = RESPONSE

$TSS = (n-1)\, SD_{DV}^2 = (124)(445.8)^2 = 24{,}643{,}467.36.$

$\sum (x_i - \bar{x})^2 = (n-1) SD_{IV}^2 = (124)(64.8)^2 = 520{,}680.96.$

$REG\ SS(IV) = r^2\, TSS = (0.91)^2\, TSS = 20{,}407{,}255.32.$

$SSE = (1 - r^2) TSS = 0.1719\, TSS = 4{,}236{,}212.04.$

ANOVA TABLE

SOURCE	DF	SS	MS	F
REG (IV)	1	20,407,255.32	20,407,255.32	592.5
ERROR	123	4,236,212.04	34,440.75	
TOTAL	124	24,643,467.36		

α	F (1, 120)	F(1, 123)	DECISION
.10	2.748	2.747	REJECT
.05	3.920	3.918	REJECT
.01	6.851	6.846	REJECT

REJECT H_0 NO LINEAR ASSOCIATION BETWEEN IV AND DV AT $\alpha = .01$ (AND $\alpha = .05$ AND $\alpha = .10$).

2C CONTINUED.

B. $\hat{\beta}_1 = r\frac{SD(DV)}{SD(IV)} = 0.91\frac{445.8}{64.8} = 6.260$

$\hat{Y}(x) = 1,667 + 6.260(x - 1,021)$

$= -4724.46 + 6.260x.$

99% CI FOR β_1:

$SE(\hat{\beta}_1) = \sqrt{\frac{MSE}{\Sigma(x_i - \bar{x}_n)^2}} = \sqrt{\frac{34,440.75}{520,680.96}} = \sqrt{0.0661}$

$= 0.257.$

99% CI FOR β_1 IS

$\hat{\beta}_1 \pm 2.617(0.257) = 6.260 \pm 0.673 = 5.587$ TO 6.933

C. $\hat{Y}(1200) = 1,667 + 6.260(1,200 - 1,021)$

$= 1,667 + 6.260(179) = 2787.54$

$\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x}_n)^2}{\Sigma(x_i - \bar{x}_n)^2}} = \sqrt{1 + \frac{1}{125} + \frac{(179)^2}{520,680.96}}$

$= \sqrt{1 + 0.008 + 0.0615} = \sqrt{1.0695} = 1.034.$

99% PME $= 2.617\sqrt{MSE}\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x}_n)^2}{\Sigma(x - \bar{x}_n)^2}}$

$= 2.617\sqrt{34,440.75}\ 1.034 = 502.27$

99% PI FOR $Y_F(1200)$ IS.

$2787.54 \pm 502.27 = 2285.27$ TO 3289.81.

X2S2022P2DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD

2. A research team conducted a longitudinal study of participants between 35 and 40 years of age. They measured each participant's average daily caloric consumption at age 35. They also measured body mass index (BMI) at age 40. BMI is a measure of obesity with higher BMI indicating more obesity. The team collected data on $n = 432$ participants. The average BMI at age 40 was 26.2, with an observed standard deviation of 5.4 (the divisor in the underlying variance calculation was $n - 1$). The average daily caloric consumption age 35 was 2,960, with an observed standard deviation of 418.8 (the divisor in the underlying variance calculation was $n - 1$). The Pearson product moment correlation coefficient between the two variables was 0.47. The research team sought to estimate the regression of BMI at age 40 on the participant's daily caloric consumption at age 35.

   a. Complete the analysis of variance table for the regression of BMI at age 40 on the participant's daily caloric consumption at age 35. Test the null hypothesis that the slope of this regression is zero at levels of significance 0.10, 0.05, and 0.01. This part is worth 25 points.

   b. Find the estimated regression of BMI at age 40 on the participant's daily caloric consumption at age 35. Find the 99% confidence interval for the slope in this equation. This part is worth 25 points.

   c. Use the least-squares equation to estimate the BMI at age 40 for participants whose daily caloric consumption at age 35 was 3,500 units. What is the 99% confidence interval for BMI at age 40 for participants whose daily caloric consumption at age 35 was 3,500 units. This part is worth 20 points.

A. IV = CALORIC CONSUMPTION DV = BMI

$$TSS = (n-1)\, SD_{DV}^2 = 431(5.4)^2 = 12{,}567.96$$

$$\Sigma(x_i - \bar{x}_m)^2 = (n-1)\, SD_{IV}^2 = (431)(418.8)^2 = 75{,}594{,}572.64.$$

$$REGSS(IV) = r^2\, TSS = (0.47)^2(12{,}567.96) = 2{,}776.26.$$

$$SSE = (1-r^2)TSS = (0.7791)\, TSS = 9{,}791.70$$

ANOVA TABLE

SOURCE	DF	SS	MS	F
REG(CONSUMPTION)	1	2,776.26	2,776.26	121.92
ERROR	430	9,791.70	22.77	
TOTAL	431	12,567.96		

α	$F(1,\infty)$	$F(1,430)$	
.10	2.706	2.717	REJECT
.05	3.841	3.863	REJECT
.01	6.635	6.694	REJECT

REJECT H_0: NO LINEAR ASSOCIATION BETWEEN CONSUMPTION AND BMI AT $\alpha = .01$ (AND $\alpha = .05$ AND $\alpha = .10$)

2D CONTINUED.

B. $\hat{\beta}_1 = r \frac{SD(DV)}{SD(IV)} \doteq (0.47) \frac{5.4}{418.8} = 0.00606$

$\hat{Y}(x) = 26.2 + 0.00606(x - 2,960)$

$\doteq 8.26 + 0.00606\, x$

99% CI FOR β_1:

$SE(\hat{\beta}_1) = \sqrt{\frac{MSE}{\Sigma(x_i - \bar{x}_m)^2}} = \sqrt{\frac{22.77}{75,594,572.64}} = 0.0005488.$

99% SM $= 2.576\ SE(\hat{\beta}_1) = 0.00141.$

99% CI FOR β_1: $0.00606 \pm 0.00141 = 0.00464$ TO 0.00747

C. $\hat{Y}(3,500) = 26.2 + 0.00606(3,500 - 2,960)$

$= 26.2 + 0.00606(540) = 29.47$

99% SM $= 2.576 \sqrt{MSE\left(\frac{1}{432} + \frac{(540)^2}{75,594,572.64}\right)}$

$= 2.576 \sqrt{22.77(0.0023148 + 0.003857)}$

$= 2.576 \sqrt{22.77(.0061721)} = 2.576 \sqrt{0.1405}$

$= 0.966$

THE 99% CI FOR $\beta_0 + 3500\beta_1$ IS

$29.47 \pm 0.966 = 28.50$ TO $30.44.$

X2S2022P345EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE

Common Information for Questions 3, 4, and 5

A research team sought to estimate the model $E(Y) = \beta_0 + \beta_1 x + \beta_2 w$. The variable Y was the participant's BMI at age 35. The variable w was the participant's daily carbohydrate consumption at age 25. The variable x was the measure of the participant's daily carbohydrate consumption at age 30. They observed values of y, x, w for $n = 576$ participants. They found that the standard deviation of Y, where the variance estimator used division by $n - 1$, was 10.2; the standard deviation of x was 281.2; and the standard deviation of w was 224.6. The correlation between Y and w was 0.43; the correlation between Y and x was 0.61; and the correlation between x and w was 0.68.

3. Compute the partial correlation coefficients $r_{Yx\bullet w}$ and $r_{Yw\bullet x}$. This question is worth 20 points.
4. Is a mediation model or an explanation model a better explanation of the observed results? This question is worth 30 points.
5. Compute the analysis of variance table for the multiple regression analysis of Y. Include the sum of squares due to the regression on w and the sum of squares due to the regression on x after including w. Test the null hypothesis that $\beta_1 = 0$ against the alternative $\beta_1 \neq 0$. Report whether the test is significant at the 0.10, 0.05, and 0.01 levels of significance. This question is worth 50 points.

End of application of common information

3. $$r_{Yx\cdot w} = \frac{0.61 - 0.43(0.68)}{\sqrt{(1-.43^2)(1-.68^2)}} = \frac{0.3176}{\sqrt{(0.8151)(0.5376)}}$$

$$= \frac{0.3176}{\sqrt{0.43820}} = \frac{0.3176}{0.66200} = 0.4794$$

$$r_{Yw\cdot x} = \frac{0.43 - 0.61(0.68)}{\sqrt{(1-.61^2)(1-.68^2)}} = \frac{0.0152}{\sqrt{(0.6279)(0.5376)}}$$

$$= \frac{0.0152}{\sqrt{0.33756}} = \frac{0.0152}{0.58100} = 0.0262$$

$$se(r_{Yw\cdot x}) \approx \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{573}} = 0.0418.$$

X IS KEY VARIABLE

4 $$\underset{25}{w} \rightarrow \underset{30}{x} \rightarrow \underset{35}{Y}$$ MEDIATION MODEL.

5 E

$$TSS = (n-1)\,SD_{DV}^2 = (576-1)(10.2)^2 = 59{,}823.0$$

$$SSREG(w) = r_{Yw}^2\,TSS = (0.43)^2\,TSS = 11{,}061.27$$

$$TSS - SSREG(w) = 48{,}761.73$$

$$SSREG(x|w) = r_{Yx\cdot w}^2\,(TSS - SSREG(w))$$

$$= (0.4798)^2\,(TSS - SSREG(w)) = 11{,}225.34$$

$$SSE = TSS - SSREG(w) - SSREG(x|w)$$

$$= 37{,}536.39$$

ANOVA TABLE

SOURCE	DF	SS	MS
REG (w)	1	11,061.27	11,061.27
REG(x|w)	1	11, 225.34	11, 225.34
ERROR	573	37, 536.39	65.51
TOTAL	575	59, 823.00	

$$F_{x|w} = \frac{SSREG(x|w)/1}{MSE} = \frac{11{,}225.34}{65.51} = 171.36$$

α	$F(1,\infty)$	$F(1,573)$	
.10	2.706	2.714	REJECT
.05	3.841	3.858	REJECT
.01	6.635	6.679	REJECT

REJECT H_0: $\beta_1 = 0$ VS H_1: $\beta_1 \neq 0$ AT $\alpha = .01$

(AND $\alpha = .05$ AND $\alpha = .10$).

X2S2022P345FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF

Common Information for Questions 3, 4, and 5

A research team sought to estimate the model $E(Y) = \beta_0 + \beta_1 x + \beta_2 w$. The variable Y was the measure of the participant's depression at age 25. The variable w was the participant's depression at age 15. The variable x was the measure of the participant's socialization at age 20. They observed values of y, x, w for $n = 457$ participants. They found that the standard deviation of Y, where the variance estimator used division by $n - 1$, was 35.6; the standard deviation of x was 43.2; and the standard deviation of w was 24.2. The correlation between Y and w was 0.71; the correlation between Y and x was ***-0.31***; and the correlation between x and w was ***-0.44***.

3 Compute the partial correlation coefficients $r_{Yx\bullet w}$ and $r_{Yw\bullet x}$. This is worth 20 points.

4 Is a mediation model or an explanation model a better explanation of the observed results? This question is worth 30 points.

5 Compute the analysis of variance table for the multiple regression analysis of Y. Include the sum of squares due to the regression on w and the sum of squares due to the regression on x after including w. Test the null hypothesis that $\beta_1 = 0$ against the alternative $\beta_1 \neq 0$. Report whether the test is significant at the 0.10, 0.05, and 0.01 levels of significance. This question is worth 50 points.

End of application of common information

3. $r_{Yx.w} = \dfrac{r_{Yx} - r_{Yw} r_{xw}}{\sqrt{(1-r^2_{Yw})(1-r^2_{xw})}} = \dfrac{-.31 - (0.71)(-.44)}{\sqrt{(1-.71^2)(1-(-.44)^2)}} = \dfrac{0.0024}{\sqrt{.4959(.8064)}}$

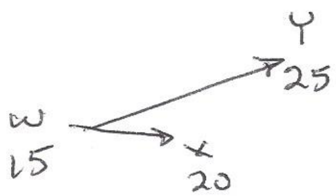
$= \dfrac{.0024}{\sqrt{.39989}} = \dfrac{.0024}{0.63237} = 0.003795.$

$\sigma(r_{Yxw}) \approx \dfrac{1}{\sqrt{n-3}} = \dfrac{1}{\sqrt{454}} = 0.0469$; W IS KEY VARIABLE

$r_{Yw.x} = \dfrac{0.71 - (-0.31)(-0.44)}{\sqrt{(1-(-0.31)^2)(1-(-0.44)^2)}} = \dfrac{0.5736}{\sqrt{(.9039)(.8064)}} = \dfrac{0.5736}{\sqrt{.72890}}$

$= \dfrac{0.5736}{0.85376} = 0.67185.$

4.



EXPLANATION.

5F. $TSS = (n-1) SD_{DV}^2 = 456 (35.6)^2 = 577,916.16$

$REGSS(w) = r_{Yw}^2 \, TSS = (0.71)^2 \, 577,916.16 = 291,327.54.$

$TSS - REGSS(w) = 286,588.62.$

$REGSS(x|w) = r_{Yx \cdot w}^2 \, (TSS - REGSS(w))$

$= (0.003795)^2 (286,588.62) = 4.13$

$SSE = TSS - REGSS(w) - REGSS(x|w) = 286,584.49$

ANOVA TABLE

SOURCE	DF	SS	MS
REG(w)	1	291,327.54	291,327.54
REG(x|w)	1	4.13	4.13
ERROR	454	286,584.49	631.24
TOTAL	456	577,916.16	

$$F_{x|w} = \frac{MS\ REG(x|w)}{MSE} = \frac{4.13}{631.24} = 0.0065.$$

α	$F(1, \infty)$	$F(1, 454)$	
.10	2.706	2.717	ACCEPT
.05	3.841	3.862	ACCEPT
.01	6.635	6.691.	ACCEPT

ACCEPT $H_0: \beta_1 = 0$ VS $H_1: \beta_1 \neq 0$ AT $\alpha = .10$

(AND $\alpha = .05$ AND $\alpha = .01$)

X2S2022P6GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG

6. A research team wishes to test the null hypothesis H_0: $\rho = 0$ at $\alpha = 0.005$ against the alternative H_1: $\rho > 0$ using Fisher's transformation of the Pearson product moment correlation coefficient as the test statistic. They have asked their consulting statistician for a sample size n such that $\beta = 0.01$ when $\rho = 0.12$. What is this value? This problem is worth 50 points.

$$\sqrt{n-3} \geq \frac{|z_\alpha| \cdot 1 + |z_\beta| \cdot 1}{(F(\rho_1) - F(0))}$$

$$F(\rho_1) = \frac{1}{2} \ln\left(\frac{1+0.12}{1-0.12}\right) = \frac{1}{2} \ln\left(\frac{1.12}{0.88}\right) = \frac{1}{2} \ln(1.272727)$$

$$= \frac{1}{2} (0.24116) = 0.12058, \quad F(\rho_0) = F(0) = 0.$$

$$\sqrt{n-3} \geq \frac{2.576(1) + 2.326(1)}{0.12058 - 0} = 40.65.$$

$$n - 3 \geq 1652.7$$

$$n \geq 1656.$$

THE SAMPLE SIZE SHOULD BE AT LEAST 1656 OBSERVATIONS.

X2S2022P6HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH

6. A research team wishes to test the null hypothesis $H_0\colon\ \rho = 0$ at $\alpha = 0.025$ against the alternative $H_1\colon\ \rho > 0$ using Fisher's transformation of the Pearson product moment correlation coefficient as the test statistic. They have asked their consulting statistician for a sample size n such that $\beta = 0.05$ when $\rho = 0.22$. What is this value? This problem is worth 50 points.

$$\rho_1 = 0.22 \quad F(\rho_1) = \frac{1}{2}\ln\left(\frac{1+0.22}{1-0.22}\right) = \frac{1}{2}\ln\left(\frac{1.22}{.78}\right)$$

$$= \frac{1}{2}\ln(1.5641) = \frac{0.44731}{2} = 0.22366.$$

$$\sqrt{n-3} \geq \frac{|z_\alpha| + |z_\beta|}{|F(\rho_1) - F(\rho_0)} = \frac{1.960 + 1.645}{0.22366 - 0} = 16.118$$

$$n - 3 \geq (16.118)^2 = 259.8$$

$$n \geq 263.$$

THE SAMPLE SIZE SHOULD BE AT LEAST 263 OBSERVATIONS.

7. The correlation matrix of the random variables Y_1, Y_2, Y_3, Y_4 is $\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$, $0 < \rho < 1$, and each random variable has variance σ^2. Let

$$W_1 = Y_1 - Y_2 - Y_3 + Y_4,$$

and let

$$W_2 = 3Y_1 - Y_2 - Y_3 + 3Y_4.$$

Find the variance covariance matrix of $(W_1, W_2)^T$. This problem is worth 50 points.

End of the Examination

$$M = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 3 & -1 & -1 & 3 \end{bmatrix}, \quad \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = M \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix}$$

$$\operatorname{vcv}\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = M \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix} M^T$$

$$= \sigma^2 \begin{bmatrix} 1-\rho & -1+\rho & -1+\rho & 1-\rho \\ 3+\rho & -1+5\rho & -1+5\rho & 3+\rho \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & -1 \\ -1 & -1 \\ 1 & 3 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 4-4\rho & 8-8\rho \\ 8-8\rho & 20-4\rho \end{bmatrix}$$

X2S2022P7JJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ

7. The correlation matrix of the random variables Y_1, Y_2, Y_3, Y_4 is $\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$, $0 < \rho < 1$, and each random variable has variance σ^2. Let

$$W_1 = Y_1 + Y_2 + Y_3 + Y_4,$$

and let

$$W_2 = 3Y_1 + Y_2 + Y_3 + 3Y_4.$$

Find the variance covariance matrix of $(W_1, W_2)^T$. This problem is worth 50 points.

End of the Examination

$\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = MY$ WHERE $M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 3 \end{bmatrix}$

$$\mathrm{vcv}\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = M \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix} M^T$$

$$= \sigma^2 \begin{bmatrix} 1+3\rho & 1+3\rho & 1+3\rho & 1+3\rho \\ 3+5\rho & 1+7\rho & 1+7\rho & 3+5\rho \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 4+12\rho & 8+24\rho \\ 8+24\rho & 20+44\rho \end{bmatrix}$$