Math 50 Final – Practice

Instructor: Ethan Levien

November 18, 2025

Name: ______________________________ Section: ________________

Instructions

- You have 3 hours to complete the exam.
- You may have a one page (single-sided) "cheat sheet" which must be turned in with the exam, but no electronics (including calculators).
- Each problem is worth 5 points.
- Write your solutions in the boxes.
- Don't cheat.

Exercise 1 (Converting code to math)**:** Consider the following code

```
x1 = np.random.normal(0,1,100)
x2 = 2*x1 + np.random.normal(0,1,100)
y = x1 + x2 + np.random.normal(0,1,100)
b1 = np.cov(y,x1)/np.var(x1)
b2 = np.cov(y,x2)/np.var(x2)
```

What are b1 and b2 (approximately)?

Solution:

Exercise 2 (Joint distribution)**:** Consider the model of a time series $X_1, X_2, X_3, \ldots$:

$$X_{i+1} | X_i \sim \text{Normal}(1 + X_i/2, 1/3)$$

(a) Write a Python function generatesim(L) to generates a simulation of L steps of this time series starting with $X_i = 0$. The function should return a length L numpy array.

(b) After many steps, the process reaches a steady state where $E[X_{i+1}] = E[X_i]$. What is the distribution of X_i in steady-state?

Exercise 3 (Regression model comparison)**:** (X, Y) data is fit to a single-predictor regression model in statsmodels using OLS, yielding the following output:

```
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          1.0685      0.186      5.756      0.000       0.700       1.437
x1             1.9622      0.177     11.062      0.000       1.610       2.314
==============================================================================
```

A second predictor X_2 is then included in the model, which yields the following output:

```
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          1.0001      0.011     92.003      0.000       0.979       1.022
x1             0.9810      0.012     82.470      0.000       0.957       1.005
x2             2.0122      0.012    168.877      0.000       1.989       2.036
==============================================================================
```

If X_2 and X_1 were fit to a linear regression model with X_2 as the response variable, what would be the regression slope?

Solution:

Exercise 4 (Sample distribution)**:** Consider the model

$$Y_1 \sim \text{Normal}(\beta_1 X_1 + \beta_2 X_2, \sigma_\epsilon^2)$$

After fitting the model, we find $\hat{\beta}_1 = 100$, $\hat{\beta}_2 = -101$, $\hat{\sigma}_\epsilon^2 = 1/4$. The model is then fit to a different data set and it is found that $\hat{\beta}_1 = -100$, $\hat{\beta}_2 = 100.4$.

(a) Is $\text{cov}(X_1, X_2)$ likely to be positive or negative for the fitted data?

Solution:

(b) Based on this information is it possible that R^2 is very close to 1?

Solution:

Exercise 5 (Interactions)**:** Consider the the two predictor linear regression model with an interaction term:

$$Y = \beta_1 X_1 + \beta_2 X_2 + J_{1,2} X_1 X_2 + \epsilon$$

The following plot shows data generated from such a model.

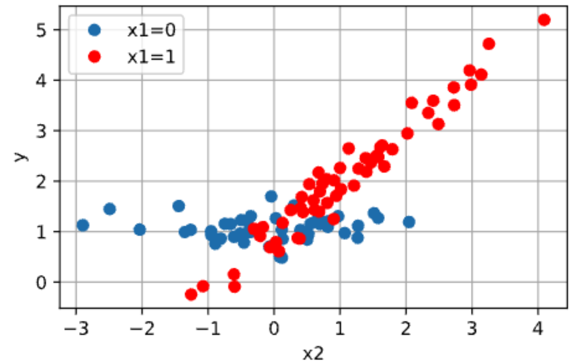

Figure 1:

What are the values of β_1, β_2 and $J_{1,2}$?

Solution:

Exercise 6 (Orthogonality)**:** Consider the features $\phi_1(x) = x^3$ and $\phi_2(x) = x^2$.

(a) Are these orthogonal with respect to $X \sim \text{Uniform}(-1, 1)$? (in the sense that $E[\phi_1(X)\phi_2(X)] = 0$)

(b) What is an example of a distribution for X such that ϕ_1 and ϕ_2 are not orthogonal?

Exercise 7 (Bayesian posterior)**:** Consider a the Bayesian model σ_ϵ and β_1, but unknown intercept with Normal priors:

$$\beta_0 \sim \text{Normal}(0, \tau_0^2) \tag{1}$$
$$Y|X, \beta \sim \text{Normal}(\beta_0 + \beta_1 X, \sigma_\epsilon^2) \tag{2}$$

Calculate the posterior of β_0.

Exercise 8 (Missing data)**:** You are given data with predictors X_1, X_2. You want to fit a linear regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

but some of the X_2 values are have been corrupted and are not reliable. One idea to handle this is called *imputation* and involves generating the missing values X_2.

Explain how you could implement imputation assuming you are a given dataframe with rows Y, X_1, X_2, C where $C = 1$ for the rows with corrupted data and $C = 0$ otherwise. What assumptions are being made for the procedure to be unbiased?

Solution:

Solutions

1. b_1 is the regression coefficient with only X_1 as a predictor and the regression coefficient $\beta_{1,2}$ of X_1 on X_2 is given to be 2 as well as $\beta_1 = \beta_2 = 1$. Thus we have $b_1 = 1+2 = 3$. The regression coefficient of X_2 on X_1 is $\beta_{2,1} = \text{cov}(X_1, X_2)/\text{var}(X_2) = \text{var}(X_1)\beta_{1,2}/(2^2+1) = 1 \times 2/5 = 2/5$. Hence $b_2 = 1 + 2/5 = 7/5$.

2. (a) One valid implementation:

```
def generatesim(L):
    X = np.zeros(L)
    for i in range(1,L):
        X[i] = 1.0 + 0.5*X[i-1] + np.random.normal(0.0, np.sqrt(1/3))
    return X
```

   (b) In steady-state we have the marginals are equal so $\mu_X = E[X_{i+1}] = 1 + E[X_i]/2 = 1 + \mu_X/2$. Thus $\mu_X = 2$. Similarly $\sigma_X^2 = \sigma_X^2/4 + 1/3$. Thus $\sigma_X^2 = 4/9$ and $X_i \sim \text{Normal}(2, 4/9)$.

3. We are given the single predictor regression coefficient $\hat{\beta}_1' \approx 2$ and the two predictor model coefficients $\hat{\beta}_1 \approx 1$, $\hat{\beta}_2 \approx 2$. Using $\hat{\beta}_1' = \hat{\beta}_1 + \hat{\beta}_2\beta_{1,2}$ we deduce the coefficient of X_1 with X_2 as the response variable, denoted $\beta_{1,2}$ here, is $\beta_{1,2} \approx 1/2$.

4. (a) This suggests the $\hat{\beta}_1$ and $\hat{\beta}_2$ are negatively correlated and therefore X_1 and X_2 are positively corrected, hence $\text{cov}(X_1, X_2) > 0$. (b) Yes, because even though we know $\hat{\sigma}_\epsilon^2$ we aren't given σ_Y^2, which could be quite large especially with such large values of $\hat{\beta}_i$.

5. β_2 is the slope of X_2 vs. Y when $X_1 = 0$, which is approximately 0 in the plot. $\beta_2 + J_{1,2}$ is the slope when $X_1 = 1$, which appears to be 1. β_1 is the expected difference in Y between $X_1 = 0$ and $X_1 = 1$ groups when $X_2 = 0$, which appears to be 0. In summary, $\beta_1 = \beta_2 = 0$ and $J_{1,2} = 1$.

6. (a) With $X \sim \text{Unif}(-1, 1)$,

$$E[\phi_1(X)\phi_2(X)] = E[X^5] = \int_{-1}^{1} x^5 \cdot \tfrac{1}{2}\, dx = 0, \tag{3}$$

   so they are orthogonal. (b) Any non-symmetric distribution makes $E[X^5] \neq 0$. For example, if $X \sim \text{Unif}(0, 1)$,

$$E[X^5] = \frac{1}{6} \neq 0, \tag{4}$$

   so they are not orthogonal.

7. The (frequentist) estimator $\hat{\beta}_0 = \overline{Y} - \beta_1\overline{X}$. This is Normal when conditioned on β_0: $\hat{\beta}_0|\beta_0 \sim \text{Normal}(\beta_0, \hat{\sigma}_\epsilon^2/N)$. If we invert this regression model we get

$$\beta_0|\hat{\beta}_0 \sim \text{Normal}\left(\beta_0\frac{\tau^2}{\tau^2 + \sigma_\epsilon^2/N}, \frac{\tau^2\sigma_\epsilon^2/N}{\tau^2 + \sigma_\epsilon^2/N}\right) \tag{5}$$

8. Using rows with $C = 0$, fit

$$X_2 = \alpha_0 + \alpha_1 X_1 + \zeta, \tag{6}$$

   and estimate $\hat{\sigma}_\zeta^2 \approx \text{var}(\zeta)$. For each $C = 1$ row,

$$\tilde{X}_2 = \hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \tilde{\zeta}, \qquad \tilde{\zeta} \sim \text{Normal}(0, \hat{\sigma}_\zeta^2). \tag{7}$$

   Then we fit

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 \tilde{X}_2 + \epsilon \tag{8}$$

   on the completed data. We have assumed C is independent of X_1 and X_2.