

**Spring 2026: Mathematical Statistics
Recitation Worksheet 13
May 1, 2026**

Linear Regression

1. Suppose $X_0 = 0$ and X_1, \dots, X_n are given by the following model, known as a Gaussian autoregressive process:

$$X_i = \theta X_{i-1} + \varepsilon_i$$

with $\theta \in [0, 1)$ unknown, and the errors $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ i.i.d. with $\sigma^2 > 0$ known.

- (a) Write down the log-likelihood function for the unknown parameter θ .
- (b) Find $\hat{\theta}_{\text{MLE}}$.
- (c) Find the least squares estimate $\tilde{\theta}$.

Hypothesis Testing

1. You collect data from one population and you want to test if the data follows a particular distribution. What test would you run in the following scenarios, or have we not covered an appropriate test?
 - (a) The data is categorical (or discrete).
 - i. You wish to test for the distribution (ie. nonparametric test).
 - ii. You wish to test for a parameter estimate.
 - (b) The data is (continuous) numerical.
 - i. You wish to test for the distribution (ie. nonparametric test).
 - ii. You wish to test for a parameter estimate.

2. You collect data from two populations and want to compare the distributions for both populations. What test would you run?
 - (a) The data is categorical and both populations can only take on two values.
 - i. Data is unpaired.
 - ii. Data is paired
 - (b) The data is continuous:
 - Data is unpaired
 - Populations follow normal distribution.
 - We want to compare the unknown distributions.
 - Data is paired
 - Populations follow normal distribution.
 - We want to compare the unknown distributions.
 - (c) The data is categorical and both populations can take on multiple values.
 - (d) What test would you run if you wanted to compare the distributions of multiple populations?