Homework 2

Due: Thursday 2/7/19 by 12:00pm (noon)

An Overview of Level-α Tests

This homework is going to ask you to conduct a level-α tests of a null hypothesis, which requires that you combine a few bits of information from class.

Let's call the null hypothesis H and the alternative hypothesis K. Suppose we have a test statistic $\hat{t}$, that is a function of the data, and that we know either exactly or approximately what the distribution of $\hat{t}$ is under the null H. Then we can construct a level-α test of the null hypothesis H by comparing $\hat{t}$ to the $1-\alpha/2$ and $\alpha/2$ quantiles of the distribution of $\hat{t}$ under the null H. If $\hat{t}$ is within those quantiles, we **accept** the null hypothesis H, otherwise we reject it.

This general idea is something that I expect have seen in your previous statistics classes, and we referenced it in our review of linear regression when we talked about testing the null hypothesis H that $\hat{\beta}_j$ is exactly equal to a specific value for some j. Consider testing the null hypothesis $H: \beta_1 = 0$ against the alternative $K: \beta_1 \neq 0$. For such a problem, our test statistic is

$$\hat{t} = \frac{\hat{\beta}_1}{s_w\sqrt{\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}_{11}}},$$

and we know that $\hat{t} \sim \mathcal{T}_{n-q}$ under H, where q is the number of covariates we have in our regression model (number of columns of $\boldsymbol{Z}$). We perform a level-α test of H by comparing $\hat{t}$ to the $1-\alpha/2$ and $\alpha/2$ quantiles of a $\mathcal{T}_{n-q}$ distribution. If $\hat{t}$ is within these quantiles, we accept $H: \beta_1 = 0$, otherwise we reject it. Here's a bit of `R` code to demonstrate what I mean, in this example:

```
# Let's work through this example with our chicken data
library(astsa)
data(chicken)
n <- length(chicken)
reg <- lm(chicken~time(chicken))
s.w <- summary(reg)$sig
q <- length(coef(reg))
# Let's test if the time coefficient is equal to zero
# First, construct the test statistic
ZtZ.inv <- solve(crossprod(model.matrix(reg)))
t.hat <- coef(reg)["time(chicken)"]/(s.w*sqrt(ZtZ.inv[2, 2]))
# Note: Another way to get this would be to set
# t.hat = summary(reg)$coef["time(chicken)", "t value"]
#
# In this case, we know that the test statistic should be t-distributed under the
# null with n-q degrees of freedom. The quantiles for a level alpha = 0.05 test will be
alpha <- 0.05
q <- qt(c(alpha/2, 1 - alpha/2), df = n - q)
# Compare the test statistic to these quantiles, do we accept?
t.hat %in% q # Accept null if TRUE, reject otherwise
```

To apply this idea to these problems, ask yourself: What is a sample quantity that we can calcuate for a time series $\boldsymbol{x}$ and use as a test statistic $\hat{t}$ that:

- We talked about in class;
- Is relevant to testing a hypothesis about $\rho_x(1)$;

- We know the approximate or exact distribution of under the null that $\boldsymbol{x}$ is a white noise time series, with $\rho_x(1) = 0$?

The AR(1) Model

1. This problem will ask you to work with the autoregressive (AR) model.
   (a) Describe what `R` returns when you run `x <- arima.sim(n = 100, list(ar=1), sd = 1)`, and why this occurs.
   (b) Simulate 1,000 $\mathbf{AR}(1)$ time series of length $n = 100$ with $\sigma_w^2 = 1$ for values of $\phi_1 = \{-0.5, -0.25, -0.125, 0, 0.125, 0.25, 0.5\}$. For each value of ϕ_1, compute the percent of simulations in which a level-0.05 test of the null hypothesis that the time series is white noise, with $\rho_x(1) = 0$, rejects the null, using $\hat{\rho}_x(1)$ from class and 3.(h) in Homework 1. Plot the percent of simulations in which a test of the null hypothesis rejects the null against ϕ_1.
   (c) When $\phi_1 \neq 0$, the percent of simulations in which a test of the null hypothesis that $\rho_x(1) = 0$ rejects the null estimates the **power** of the test. When $\phi_1 = 0$, the percent of simulations in which a test of the null hypothesis that $\rho_x(1) = 0$ rejects the null estimates the **level** of the test. Is the estimated level 0.05, as we would expect from a level-0.05 test? If not, why not?
   (d) Describe in at most two sentences how the power of the test relates to the true value ϕ_1. Intuitively, does this make sense?

The MA(1) Model

2. This problem will ask you to work with the moving average (MA) model.
   (a) Without using the `arima.sim` function or any other third party function for simulating an MA time series, simulate a length-100 time series $\boldsymbol{x}$ according to the MA model:

$$x_t = 0.5w_{t-1} + w_t,\ w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$$

   (b) Using the code you wrote in (a) or `arima.sim`, simulate 1,000 $\mathbf{MA}(1)$ time series of length $n = 100$ with $\sigma_w^2 = 1$ for values of $\theta_1 = \{-1, -0.268, -0.127, 0, 0.127, 0.268, 1\}$. For each value of θ_1, compute the percent of simulations in which a test of the null hypothesis that the time series is white noise, with $\rho_x(1) = 0$, rejects the null, using $\hat{\rho}_x(1)$ from class and 3.(h) in Homework 1. Plot the percent of simulations in which a test of the null hypothesis rejects the null against θ_1.
   (c) When $\theta_1 \neq 0$, the percent of simulations in which a test of the null hypothesis that $\rho_x(1) = 0$ rejects the null estimates the **power** of the test. When $\theta_1 = 0$, the percent of simulations in which a test of the null hypothesis that $\rho_x(1) = 0$ rejects the null estimates the **level** of the test. Is the estimated level 0.05? If not, why not?
   (d) Describe in at most two sentences how the power of the test relates to the true value θ_1. Intuitively, does this make sense?

Comparing AR(1) and MA(1) Models

3. This problem asks you to compare what you observed in 1. (b)-(d) to what you observed in 2. (b)-(d).
   (a) Compute the true lag-one autocorrelation $\rho_x(1)$ under for an $\mathbf{AR}(1)$ model with $\phi_1 = \{-0.5, -0.25, -0.125, 0.125, 0.25, 0.5\}$.
   (b) Compute the true lag-one autocorrelation $\rho_x(1)$ under for an $\mathbf{MA}(1)$ model with $\theta_1 = \{-1, -0.268, -0.127, 0, 0.127, 0.268, 1\}$.
   (c) Plot the the percent of simulations in which a test of the null hypothesis rejects the null against the true autocorrelation $\rho_x(1)$ for both the $\mathbf{AR}(1)$ and $\mathbf{MA}(1)$ simulations on a single plot. You

should have two lines or sets of points, one for the **AR**(1) simulations and one for the **MA**(1) simulations.

(d) In one sentence, interpret what you observe in (c), taking what you find in (a) and (b) into account.