Homework for Econ 312 Chapter 6

May 5, 2026

1. Consider the same i.i.d. Gaussian setup as in the previous question, but now treat $\theta = \mu$ as an unknown scalar parameter and set $\mu_0 = 0$, $\sigma^2 = 1$ as the baseline. The single-observation log-likelihood is
$$\log \psi(z \mid \mu) = -\frac{1}{2}(z-\mu)^2 - \frac{1}{2}\log(2\pi).$$
   (a) Compute the score increment $\frac{d}{d\mu}\log\psi(Z_t \mid \mu_0)$ and verify that its expectation under the true model $\mu_0 = 0$ equals zero, confirming the information equation.

   (b) Compute the Fisher information
$$V = E\left[\left(\frac{d}{d\mu}\log\psi(Z_t \mid 0)\right)^2\right]$$
directly from the score, then verify the **information equality** by computing
$$V = -E_0\left[\frac{d^2}{d\mu^2}\log\psi(Z_t \mid 0)\right]$$
and confirming that both expressions give the same value.

   (c) Now add a nuisance parameter: suppose the variance σ^2 is also unknown, with true value $\sigma_0^2 = 1$. Let $\theta = \mu$ be the parameter of interest and $\vartheta = \sigma^2$ the nuisance parameter. Compute the 2×2 Fisher information matrix $\mathbb{V}$ for (θ, ϑ), obtain the population regression coefficient β from projecting the score for μ on the score for σ^2, and compute the residual variance $E(U_{t+1}^2)$. Verify the nuisance parameter inequality and interpret the result.

2. Consider two possible models:
$$Z_{t+1} = \mathsf{n}_0 + (\mathsf{d}_0)' X_t + \mathsf{f}_0 W_{t+1}$$
$$Z_{t+1} = \mathsf{n}_1 + (\mathsf{d}_1)' X_t + \mathsf{f}_1 W_{t+1}.$$

where Z_{t+1} a scalar that is observed for a finite sample, W_{t+1} is a scalar, standard normally distributed shock, and X_t is vector consisting of Z_t and a finite number of lags. Each model specifies the unknown coefficients $\theta = (\mathsf{n}, \mathsf{d}, \mathsf{f})$ where $\theta_1 \neq \theta_0$.

(a) Use the date $t+1$ contribution to the log-likelihood ratio of model 1 relative to model 2 to form:

$$\kappa(X_t, W_{t+1})$$

(b) Suppose that model θ_0 is what generates the data. Verify:

i.

$$E\left(\exp[\kappa(X_t, W_{t+1})] \mid X_t, \theta_0\right) = 1,$$

which implies that the multiplicative functional is a martingale.

ii.

$$E\left(\kappa(X_t, W_{t+1}) \mid X_t, \theta_0\right) < 0$$

which implies that the log-likelihood increment is expected to decline over time as data accumulates.

iii. For $0 < \alpha < 1$,

$$E\left(\exp[\alpha\kappa(X_t, W_{t+1})] \mid X_t, \theta_0\right) < 0,$$

which is used in the study of fixed threshold rules for model selection.

In checking these relations, you may use the following result. If W is has a standard normal distribution then

$$E\left[\exp\left(\mathsf{a} + \mathsf{b}W + \frac{\mathsf{c}}{2}W^2\right)\right] = \frac{1}{\sqrt{1-\mathsf{c}}}\exp\left[\mathsf{a} + \frac{\mathsf{b}^2}{2(1-\mathsf{c})}\right]$$

where $\mathsf{c} < 1$. You **do not** have to derive this formula, although it can be shown with a "complete-the-square" argument.

3. Consider the setup in the previous question. We next illustrate the computation of Chernoff entropy. While Chapter 8 provides general methods for doing this, in this question we add a simplification to make the computation particularly easy. Suppose $\mathsf{d}_1 = \mathsf{d}_2 = 0.$. Use your formulas for:

$$E\left(\exp[\alpha\kappa(X_t, W_{t+1})] \mid X_t, \theta_0\right)$$

except that X_t drops out of the formula.

(a) Form the negative of the logarithm as a function of α and maximize over $0 < \alpha < 1$ as is required for Chernov entropy. You may preform the maximization either analytically or graphically for a few different parameter configurations.

(b) Discuss what the impact is of increasing f_0 and f_1 holding fixed n_0 and n_1. Also discuss the impact of increasing the difference between n_0 and n_1 holding fixed $\mathsf{f}_0 = \mathsf{f}_1$. You may do this analytically or numerically.