Name   David Plazas Escudero

Student code   201710005101

1. *Estimate the expected value of an AR(1) model with constant term, i.e. $x_t = \mu + \psi x_{t-1} + \varepsilon_t$.*

   Let $\{x_t\}_{t \geq 0}$ be an stochastic process driven by an AR(1) model $x_t = \mu + \psi x_{t-1} + \varepsilon_t$, with $\mu \in \mathbb{R}$, with ψ such that $|\psi| < 1$ (for stationarity) and $\{\varepsilon_t\}_{t \geq 0}$ is a white-noise process. Let us calculate the expected value.

   Taking the expected value to both sides we have

   $$\mathbb{E}\left[x_t\right] = \mathbb{E}\left[\mu + \psi x_{t-1} + \varepsilon_t\right].$$

   Recall that the mathematical expectation $\mathbb{E}\left[\cdot\right]$ is a linear operator, hence

   $$\mathbb{E}\left[x_t\right] = \mathbb{E}\left[\mu\right] + \psi\mathbb{E}\left[x_{t-1}\right] + \mathbb{E}\left[\varepsilon_t\right].$$

   It is clear that the expectation of a non-stochastic term is the term itself and, given that $\{\varepsilon_t\}_{t \geq 0}$ is a white-noise process, it has a expected value of exactly 0. Therefore,

   $$\mathbb{E}\left[x_t\right] = \mu + \psi\mathbb{E}\left[x_{t-1}\right]$$

   and, as we are working under the assumption of stationarity, the expected value should remain the same over time, formally $\mathbb{E}\left[x_t\right] = \mathbb{E}\left[x_{t-s}\right]$, $\forall s \geq 0$. Thus,

   $$\mathbb{E}\left[x_t\right] = \mu + \psi\mathbb{E}\left[x_t\right]$$

   and solving for $\mathbb{E}\left[x_t\right]$ we have the desired expression

   $$\mathbb{E}\left[x_t\right] = \frac{\mu}{1 - \psi}.$$

2. *Under the assumptions studied in class, show that $\mathbb{E}\left[\varepsilon_t \varepsilon_{t-1}\right] = 0$.*

   Let $\{\varepsilon_t\}_{t \geq 0}$ be a white-noise process. Recall the alternative formula for the covariance of two random variables X and Y:

   $$\mathrm{Cov}\left[X, Y\right] = \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right].$$

   In our specific case $X \equiv \varepsilon_t$ and $Y \equiv \varepsilon_{t-1}$, hence

   $$\mathrm{Cov}\left[\varepsilon_t, \varepsilon_{t-1}\right] = \mathbb{E}\left[\varepsilon_t \varepsilon_{t-1}\right] - \mathbb{E}\left[\varepsilon_t\right]\mathbb{E}\left[\varepsilon_{t-1}\right]$$

   and we can solve for the desired term as follows:

   $$\mathbb{E}\left[\varepsilon_t \varepsilon_{t-1}\right] = \mathrm{Cov}\left[\varepsilon_t, \varepsilon_{t-1}\right] + \mathbb{E}\left[\varepsilon_t\right]\mathbb{E}\left[\varepsilon_{t-1}\right].$$

   Recall the assumption of $\{\varepsilon_t\}_{t \geq 0}$ being a white-noise process, then $\mathbb{E}\left[\varepsilon_t\right] = 0$, $\forall t \geq 0$; and the process is not autocorrelated, i.e. $\mathrm{Cov}\left[\varepsilon_t, \varepsilon_{t-s}\right] = 0$, $\forall s \geq 0$. Replacing we obtain

   $$\mathbb{E}\left[\varepsilon_t \varepsilon_{t-1}\right] = 0.$$

3. *Consider the model $x_t = 1 + 0.6x_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0,2)$. Generate two series, one with 100 observations and the other one with 5000. Which mean and variance are closer to the theoretical value? Why is there a difference between these series?*

Let us present first the code in Python to simulate the model at hand:

```python
import numpy as np
import scipy.stats as st

# Parameters
mu = 1
psi = 0.6
sigma2 = 2

def ar1(T):
    e = st.norm.rvs(size=(T), loc=0, scale=sigma2**0.5)
    xs = np.zeros(T)
    xs[0] = mu
    for i in range(1, len(xs)):
        xs[i] = mu + psi * xs[i-1] + e[i]
    return xs

xs1 = ar1(100)
xs2 = ar1(5000)
mean1, var1 = xs1.mean(), xs1.var(ddof=1)
mean2, var2 = xs2.mean(), xs2.var(ddof=1)
```

This AR(1) model has parameters $\mu = 1$, $\psi = 0.6$ and $\sigma^2 = 2$. First, the theoretical expected value is $\mathbb{E}[x_t] = 2.5$ and the results for 1 experiment in both series are 2.289 and 2.551 respectively. For the variance, the theoretical value is $\mathbb{V}[x_t] = 3.125$, and the estimated values were 2.925 and 3.146 respectively.

As expected, the second experiment, the one with 5000 observations, is closer to the theoretical values of the expected value and variance. There is a significant difference between these values between the two series, since the sample sizes differ and it is clear, by the law of large numbers, that as the number of experiments is increased, the estimated results should converge on the true values.

The End.