Solution to Exercise 2

Simon A. Broda

1. (a) The first thing to observe is the difference between the population quantities (parameters) μ and σ, and their estimates $\bar{y}$ and s_y, which are sample quantities. The latter will generally be close to the former because of the law of large numbers, but not the same. The estimates also change every time new random numbers are drawn. The time series plot shows that the observations are randomly scattered around 0. The autocorrelations in the correlogram will mostly be insignificant, i.e., statistically indistinguishable from zero. Formally, the hypothesis being tested is $H_0 : \tau_s = 0$ vs. $H_a : \tau_s \neq 0$. Even if H_0 is true here, there is still a 5% probability that any given autocorrelation will be significant; this is called the type-1 error: the probability of rejecting the null even though it is true. A similar statement concerns the Q-statistics, which test whether the first m autocorrelations are all equal to zero[1]. **Important**: make sure to look at the formulas behind the cells and make sure you understand how they work; specifically, the correlogram, the Q-stats, and their respective critical values. You don't need to understand how the random numbers u_t themselves are generated (they use a trick called inverse transform sampling).

   (b) The time series plot looks very different from that in the other sheet, because a random walk is not mean reverting. Also, the correlogram and the Q-stats are now highly significant, so that we (correctly) reject the null that the data were generated from a white noise process. In fact, the slow and almost linear decay of the correlogram suggests (correctly) that the data were generated by an integrated process. **Important**: make sure you understand how the simulated random walk y_t is constructed, by always taking "yesterday's" value y_{t-1} and adding u_t to it.

2. See Jupyter notebook.

3. (a) By repeatedly plugging in,

$$
\begin{aligned}
Y_t &= Y_{t-1} + U_t \\
&= Y_{t-2} + U_{t-1} + U_t \\
&= Y_{t-3} + U_{t-2} + U_{t-1} + U_t \\
&\vdots \\
&= Y_0 + \sum_{s=1}^{t} U_s
\end{aligned}
$$

   as claimed.

[1]Formally, $Q(m)$ can be used to test $H_0 : \tau_1 = \tau_2 = \ldots = \tau_m = 0$.

(b) The result from the previous question implies that

$$\begin{aligned}\mathbb{E}\left[Y_t\right] &= \mathbb{E}\left[Y_0 + \sum_{s=1}^{t} U_s\right] \\ &= Y_0 + \sum_{s=1}^{t} \mathbb{E}[U_s] = Y_0.\end{aligned}$$

Here, we have used that the expectation of a sum is the sum of the expectations, together with the fact that Y_0 is assumed to be a constant, and that $\mathbb{E}[U_t] = 0$ because U_t is white noise. For the variance, recall that in general, $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\,\mathrm{cov}(X,Y)$. But since the U_t are all independent, we have that $\mathrm{var}(U_s + U_t) = \mathrm{var}(U_s) + \mathrm{var}(U_t) + 0 = 2\sigma^2$. Thus

$$\begin{aligned}\mathrm{var}\left[Y_t\right] &= \mathrm{var}\left[Y_0 + \sum_{s=1}^{t} U_s\right] \\ &= \mathrm{var}\left[\sum_{s=1}^{t} U_s\right] \\ &= \sum_{s=1}^{t} \mathrm{var}(U_s) \\ &= \sum_{s=1}^{t} \sigma^2 \\ &= t \cdot \sigma^2.\end{aligned}$$