

Chapter 1 lecture questions

Q1: What is $E[x]$ for a fair die?

Answer: 3.5

Solution:

For a fair die, $P_i = 1/6$, so

$$E[x] = \sum_{i=1}^6 x_i P_i = (1 + 2 + 3 + 4 + 5 + 6)/6 = 21/6 = 3.5.$$

Q2: Suppose for a time series with 500 monthly values the autocorrelation function has values 0.72, 0.50, 0.31, 0.18, 0.04, -0.21, -0.33, -0.45, -0.24, -0.11, 0.15, 0.29, 0.37, 0.22, 0.06, ... for lags of 1, 2, ..., 15. Estimate the first zero crossing l_0 and the effective sample size.

Answer: $l_0 \approx 5$ and the effective sample size $N_{\text{eff}} \approx 100$.

Solution:

As the lag increases, the autocorrelation function goes from being positive to negative then to positive again ... At lag = 5 months, the autocorrelation has dropped to 0.04, so it becomes zero just beyond lag = 5. Taking the first zero crossing $l_0 \approx 5$ months, the effective sample size $N_{\text{eff}} = N/l_0 = 500/5 = 100$.

Q3: If $\rho_{xy} = 0.7$, what % of the variance of y is not accounted for by the regression?

Answer: 51%.

Solution:

If $\rho_{xy} = 0.7$, then $1 - \rho_{xy}^2 = 1 - 0.49 = 0.51$, i.e. 51% of the variance of y is not accounted for by the regression. Note that although 0.7 looks like a reasonably high value for the correlation, only 49% of the variance is explained by the linear regression, with the remainder 51% not accounted for.
