EXERCISE ON OLS, PROBABILITY INTERVALS, ASYMPTOTICS

(1) For each of the $\{X_n\}$, Z pairs below, determine whether $X_n \xrightarrow[n\to\infty]{Q} Z$, where Q takes on the values P, D, $q.m.$, and $a.s.$. Explain your answers. Note that in some of the cases where there is convergence in probability, determining whether there is a.s. convergence may be challenging.
  You can use the central limit theorem and the strong law of large numbers.
  (a) $X_n = Z + Y_n$, where Y_n is $N(0, 1/n)$ for each n and independent across n.
  (b) X_n independent across n, $P[X_n = 1/n] = 1 - 1/n$, $P[X_n = \sqrt{n}] = 1/n$, $P[Z = 0] = 1$.
  (c) X_n is the X_n from 1b, squared, $P[Z = 0] = 1$
  (d) $X_n = \sum_1^n W_j/n + 1/n$, where $\{W_j\}$ is an i.i.d. sequence with finite variance and $E[W_j] = \mu$, $P[Z = \mu] = 1$
  (e) $X_n = \sum_1^n W_j/\sqrt{n + \sqrt{n}}$, $\{W_j\}$ i.i.d. with mean zero and variance 1, $Z \sim N(0, 1)$.

(2) Compute a linear regression of `testscr` on `str`, `avginc`, `meal_pct` `el_pct` `teachers` and `computer`, using the `caschool` data set.
  (a) Form a χ-squared or F statistic to test the hypothesis that the coefficients on `teachers` and `computer` are both zero. Comment on whether it suggests these coefficients as a pair are unimportant in explaining `testscr`. [In R, the easiest way to do this is to get the package `car` (companion to applied regression), which you should be able to get with `install.packages(car)`. You then load the package with `library(car)`. After that several useful commands are available, including `linearHypothesis()`, which provides an easy way to generate test statistics like the one you are asked for here. With slightly more effort, you can do without `car` and use the function `vcov(lmout)` (to get the coefficient covariance matrix), then a matrix expression to get the chi-squared or F statistic, then `pchisq()` or `pf()`.]
  (b) Find the estimated correlation matrix of the coefficients on these two variables, and use it to sketch a few representative level curves of their flat-prior joint posterior pdf in $\mathbb{R}^2$. [In R, `vcov(lmout)` delivers the coefficient covariance matrix for a linear regression, and `cov2cor(W)` converts the covariance matrix `W` to a correlation matrix. The `car` package has a function that computes and plots confidence ellipsoids (which are also flat-prior probability ellipsoids) from the output of `lm`.]
  (c) Both `teachers` and `computer` are simple counts over the whole school, so they vary with the size of the school. If the size of the school were important, one might expect that the two variables would enter with the same sign. Another possibility is that it is computers *per teacher* that matters, so that we would expect the two coefficients to have opposite signs. Generate a sample of 10000 draws from the joint posterior distribution on β, σ^2 (the regression coefficients and the residual variance) under a flat prior on $\log \sigma^2$ and β. Use it to assess the probability that the coefficients on `teachers` and `computer` are:
    (i) both positive,
    (ii) both negative,
    (iii) positive on `teachers` and negative on `computer`
    (iv) positive on `computer` and negative on `teachers`.
  What do these results imply about the evidence on whether there are school size effects?