StatsI — Exercise 4

1. (a) Let X be a random variable with mean μ and variance $\sigma^2 \in (0, \infty)$. Show that $P(|X - \mu| > 2\sigma) \leq 0.25$. What does this inequality tell us about the distribution of X?

   (b) Let $X_1, \cdots, X_n$ be an IID sample from a population with mean μ and variance $\sigma^2 \in (0, \infty)$. Show that for any $\epsilon > 0$, $P(|\bar{X}_n - \mu| > \epsilon\sigma) \leq \frac{1}{n\epsilon^2}$. Compare this bound with the approximation implied by the CLT when n is large.

   Note. The conditions required for these inequalities are minimum.

2. Let X and Y be two r.v.s with positive and finite variances. The correlation coefficient of X and Y is defined as $\rho = \mathrm{Cov}(X, Y) / \sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$. (If $\rho = 0$, X and Y are called uncorrelated or linearly independent.)

   (a) Show that $|\rho| \leq 1$.

   (b) If $Y = aX + b$ for some constants $a (\neq 0)$ and b, show that $|\rho| = 1$.

   Note. In fact $|\rho| = 1$ if and only if $Y = aX + b$ for some constants $a (\neq 0)$ and b.

3. Let $X_1, \cdots, X_n$ be a sample from a distribution with mean μ and variance $\sigma^2 \in (0, 1)$. Let $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$, where $\bar{X}_n$ is the sample mean.

   (a) Show that $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} X_n^2 - \frac{n}{n-1} \bar{X}_n^2$.

   (b) Using Slutsky's theorem, show that $S_n^2 \xrightarrow{P} \sigma^2$.