Практикум 7. Нормальная модель.

1.1. С.в. $X_n \sim \chi^2_n$. Построить графики плотности с.в. X_n, X_n/n, $(X_n - n)/\sqrt{2n}$ для $n = 1, 2, 3, 4, 5, 10, 20, 50, 100$.
1.2. Пусть $X_1, X_2 \ldots$ – н.о.р. $S_n = X_1 + \cdots + X_n$. Согласно ЦПТ распределение с.в. $Y_n = (S_n - \mathbf{E}S_n)/\sqrt{\mathbf{D}S - n}$ стремится к $\mathcal{N}(0,1)$. Сгенерировать выборку с.в. $Y_{n,1}, \ldots, Y_{n,m}$. Сравнить распределение с.в. $\sum_{i=1}^m Y_{n,i}^2$ с распределением χ^2_m, если $n = 1, 2, 5, 10, 50$ и а) $X_i \sim R[0,1]$, а) $X_i \sim Bern(0.1)$, в) X_i – смесь $\mathcal{N}(0,1)$ и $\mathcal{N}(0,10)$ с весами $9/10$, $1/10$ (описание смеси см. в практикуме 4).
1.3. $X_1, \ldots, X_n \sim \mathcal{N}(a, \sigma^2)$, a неизвестно. Доверительный интервал для σ^2 можно построить, используя тот факт, что $nS^2/\sigma^2 \sim \chi^2_{n-1}$, где $nS^2 = \sum_{i=1}^n (X_i - \overline{X})^2$. Если не слишком внимательно слушать лекции, то можно предположить, что nS^2/σ^2 имеет распределение χ^2_n. Сравнить вероятности попадания σ^2 в правильный и неправильный доверительный интервалы. Каждый интервал имеет вид

$$\sigma^2 \in \left(\frac{nS^2}{x_{1-\alpha/2}}, \frac{nS^2}{x_{\alpha/2}} \right),$$

где x_α – квантиль распределения χ^2_{n-1} для правильного и χ^2_n для правильного интервалов.

2.1. Построить графики плотности распределения Стьюдента с n степенями свободы для $n = 1, 2, 5, 10, 50$. Сравнить со стандартной нормальной плотностью. Для $n = 1$ сравнить с плотностью распредления Коши.
2.2. Пусть $X_1, X_2 \ldots$ – н.о.р. $S_n = X_1 + \cdots + X_n$. Согласно ЦПТ распределение с.в. $Y_n = (S_n - \mathbf{E}S_n)/\sqrt{\mathbf{D}S - n}$ стремится к $\mathcal{N}(0,1)$. Сгенерировать выборку с.в. $Y_{n,1}, \ldots, Y_{n,m+1}$. Сравнить распределение с.в. $Y_{n,1}/\sqrt{\sum_{i=2}^{m+1} Y_{n,i}^2/m}$ с распределением Стьюдента с m степенями свободы, если $n = 1, 2, 5, 10, 50$ и а) $X_i \sim R[0,1]$, а) $X_i \sim Bern(0.1)$, в) X_i – смесь $\mathcal{N}(0,1)$ и $\mathcal{N}(0,10)$ с весами $9/10$, $1/10$ (описание смеси см. в практикуме 4).
2.3. $X_1, \ldots, X_n \sim \mathcal{N}(a, \sigma^2)$, σ неизвестно. Доверительный интервал для a можно построить методом подстановки оценки неизвестной дисперсии, а можно использовать распределение Стьюдента. Сравнить вероятности попадания в эти интервалы при разных n. Границы каждого интервала имеют вид

$$\overline{X} \pm x_{1-\alpha/2} \sqrt{\sum_{i=1}^n (X_i - \overline{X})^2/(n-1)},$$

где x_α – квантиль распределения $\mathcal{N}(0,1)$ для первого интервала, и распределения Стьюдента с $n-1$ степенями свободы для второго интервала.

3.1. Построить графики плотности распределения Фишера для разных значений параметров m, n. Как себя ведут эти плотности при а) фиксированном m, $n \to \infty$, б) $m \to \infty$, n фиксировано, в) $m, n \to \infty$?
3.2. Пусть $X_1, X_2 \ldots$ – н.о.р. $S_n = X_1 + \cdots + X_n$. Согласно ЦПТ распределение с.в. $Y_n = (S_n - \mathbf{E}S_n)/\sqrt{\mathbf{D}S - n}$ стремится к $\mathcal{N}(0,1)$. Сгенерировать выборку с.в. $Y_{n,1}, \ldots, Y_{n,m+k}$. Сравнить распределение с.в.

$$\frac{\sum_{i=1}^m Y_{n,i}^2/m}{\sum_{i=m+1}^{m+k} Y_{n,i}^2/k}$$

с распределением Фишера с параметрами m, k, если $n = 1, 2, 5, 10, 50$ и а) $X_i \sim R[0,1]$, а) $X_i \sim Bern(0.1)$, в) X_i – смесь $\mathcal{N}(0,1)$ и $\mathcal{N}(0,10)$ с весами $9/10$, $1/10$ (описание смеси см. в практикуме 4).
3.3. $X_1, \ldots, X_n \sim \mathcal{N}(a_1, \sigma_1^2)$, $Y_1, \ldots, Y_m \sim \mathcal{N}(a_2, \sigma_2^2)$ a_1, a_2 неизвестны. Доверительный интервал для σ_2^2/σ_1^2 можно построить, используя тот факт, что

$$\frac{nS_X^2/(\sigma_1^2(n-1))}{mS_Y^2/(\sigma_2^2(m-1))} \sim F_{n-1,m-1},$$

где $nS_X^2 = \sum_{i=1}^n (X_i - \overline{X})^2$, $nS_Y^2 = \sum_{i=1}^m (Y_i - \overline{Y})^2$. Доверительный интервал имеет вид

$$\frac{\sigma_2^2}{\sigma_1^2} \in \left(f_{1-\alpha/2} \frac{mS_Y^2/(m-1)}{nS_X^2/(n-1)}, f_{\alpha/2} \frac{mS_Y^2/(m-1)}{nS_X^2/(n-1)} \right),$$

$f_{\alpha/2}$ – квантили распределения Фишера $F_{n-1,m-1}$.

Если не слишком внимательно слушать лекции, то можно предположить, что nS^2/σ^2 имеет распределение χ^2_n. Сравнить вероятности попадания σ_2^2/σ_1^2 в правильный и неправильный доверительный интервалы. Неправильный интервал имеет вид

$$\frac{\sigma_2^2}{\sigma_1^2} \in \left(f_{1-\alpha/2} \frac{S_Y^2}{S_X^2}, f_{\alpha/2} \frac{S_Y^2}{S_X^2} \right),$$

$f_{\alpha/2}$ – квантили распределения Фишера $F_{n,m}$.