1. Имеется выборка объема 3 из нормального распределения с неизвестными математическим ожиданием и дисперсией. Несмещенная выборочная дисперсия равна 1. Не пользуясь таблицами, построить точный доверительный интервал для неизвестной дисперсии уровня доверия 0.9.

Указание. $\ln 0.05 \approx -3$; $\ln 0.95 \approx -0.05$.

Решение. Точный доверительный интервал для дисперсии уровня доверия 0.9 в данном случае имеет вид

$$\left[\frac{2}{\tau_{0.95}}; \frac{2}{\tau_{0.05}}\right],$$

где τ_δ — квантиль уровня δ распределения H_2 (хи-квадрат с двумя степенями свободы). Но $H_2 \equiv E_{1/2}$, поэтому $\mathsf{P}(\chi_2^2 < x) = 1 - e^{-x/2}$ при $x > 0$. Тогда τ_δ удовлетворяет уравнению $1 - e^{-\tau_\delta/2} = \delta$, или $\tau_\delta = -2\ln(1-\delta)$.

Следовательно, $\tau_{0.95} = -2\ln 0.05 \approx 6$, $\tau_{0.05} = -2\ln 0.95 \approx 0.1$, и интервал имеет вид $[\frac{1}{3}; 20]$.

2. Проверяется гипотеза о близости математических ожиданий двух независимых нормальных выборок объемов n и m с единичными дисперсиями: $X_1, \ldots, X_n \in \mathbf{N}_{a_1,1}$, $Y_1, \ldots, Y_m \in \mathbf{N}_{a_2,1}$. Для проверки основной гипотезы H_1: $a_1 = a_2$ против альтернативы H_2: $a_1 - a_2 > 2$ используется критерий δ, принимающий основную гипотезу, если $\overline{X} - \overline{Y} \leqslant 2$, и альтернативу — в противном случае. Проверить, является ли этот критерий состоятельным (при $n, m \to \infty$). Подробно обосновать все предельные переходы.

Решение. Вероятность ошибки второго рода предложенного критерия равна

$$\mathsf{P}_{H_2}(\overline{X} - \overline{Y} \leqslant 2).$$

При верной H_2, согласно ЗБЧ, $\overline{X} - \overline{Y} \xrightarrow{\text{p}} a_1 - a_2 > 2$. Сходимость по вероятности к постоянной эквивалентна слабой сходимости, поэтому $F_{\overline{X}-\overline{Y}}(x) \to F_{a_1-a_2}(x)$ для всех $x \neq a_1 - a_2$ (точек непрерывности ф. р. $F_{a_1-a_2}(x)$). Но $2 \neq a_1 - a_2$ при верной H_2, так что $F_{\overline{X}-\overline{Y}}(2) \to F_{a_1-a_2}(2)$, и

$$\mathsf{P}_{H_2}(\overline{X} - \overline{Y} \leqslant 2) = \mathsf{P}_{H_2}(\overline{X} - \overline{Y} < 2) \to \mathsf{P}_{H_2}(a_1 - a_2 < 2) = \mathsf{P}_{H_2}(a_1 - a_2 \leqslant 2) = 0,$$

нестрогое неравенство здесь ничего не меняет, так как обе функции распределения непрерывны в точке 2.

Вспомните, что сходимость $\mathsf{P}(\overline{X} < c) \to \mathsf{P}(\mathsf{E}\, X_1 < c)$ *имеет место не при всех* c.

3. В условиях задачи **2** при любых фиксированных n и m найти вероятность ошибки первого рода критерия δ (выразить ее в терминах функции распределения стандартного нормального закона).

Решение. Вероятность ошибки первого рода предложенного критерия равна

$$\mathsf{P}_{H_1}(\overline{X} - \overline{Y} > 2).$$

При верной H_1 величина $\overline{X} - \overline{Y}$ имеет нормальное распределение с параметрами 0 и $\frac{1}{n} + \frac{1}{m} = \frac{m+n}{mn}$. Поэтому

$$\mathsf{P}_{H_1}(\overline{X} - \overline{Y} > 2) = \mathsf{P}_{H_1}\left((\overline{X} - \overline{Y}) \cdot \sqrt{\frac{mn}{m+n}} > 2\sqrt{\frac{mn}{m+n}}\right) = 1 - \Phi_{0,1}\left(2\sqrt{\frac{mn}{m+n}}\right).$$

4. Дана выборка $X_1, \ldots, X_n$ объема n. Рассматриваются две простые гипотезы: H_1: X_i имеют распределение с плотностью f_1 и H_2: X_i имеют распределение с плотностью f_2. Здесь $f_1(y) = 4\mathrm{e}^{-4(y-2)} I(y \geqslant 2)$; $\quad f_2(y) = 2\mathrm{e}^{-2(y-1)} I(y \geqslant 1)$.

Найти пределы при $n \to \infty$ вероятностей ошибок первого и второго рода следующих критериев:

$$\delta_1 = \begin{cases} H_1, & \text{если } \overline{X} \geqslant 2.25 - 1/(8\sqrt{n}), \\ H_2, & \text{если } \overline{X} < 2.25 - 1/(8\sqrt{n}); \end{cases} \quad \text{и} \quad \delta_2 = \begin{cases} H_1, & \text{если } \overline{X} \geqslant 2.25, \\ H_2, & \text{если } \overline{X} < 2.25. \end{cases}$$

Указание. Для вычисления моментов полезно вспомнить, как меняется плотность распределения при линейном преобразовании случайной величины. $\Phi_{0,1}(1) \approx 0.84$; $\Phi_{0,1}(2) \approx 0.97$; $\Phi_{0,1}(0.5) \approx 0.7$.

Решение.
Пусть $\xi \Subset E_4, \eta \Subset E_2$. Тогда f_1 — плотность распределения величины $\xi + 2$, f_2 — величины $\eta + 1$. То есть

$$\mathsf{E}_{H_1} X_1 = \mathsf{E}(\xi + 2) = 2.25,\ \mathsf{D}_{H_1} X_1 = \mathsf{D}(\xi + 2) = \frac{1}{16},\ \mathsf{E}_{H_2} X_1 = \mathsf{E}(\eta + 1) = 1.5,\ \mathsf{D}_{H_2} X_1 = \mathsf{D}(\eta + 1) = \frac{1}{4}.$$

Для первого критерия

$$\alpha_1(\delta_1) = \mathsf{P}_{H_1}(\overline{X} < 2.25 - 1/(8\sqrt{n})) = \mathsf{P}_{H_1}(\sqrt{n \cdot 16}(\overline{X} - 2.25) < -1/2) \to \Phi_{0,1}(-1/2) \approx 0.3.$$

$$\alpha_2(\delta_1) = \mathsf{P}_{H_2}(\overline{X} \geqslant 2.25 - 1/(8\sqrt{n})) = \mathsf{P}_{H_2}(\overline{X} + 1/(8\sqrt{n}) \geqslant 2.25) \to \mathsf{P}(1.5 \geqslant 2.25) = 0.$$

Предельный переход в последней формуле возможен, так как по ЗБЧ $\overline{X} + 1/(8\sqrt{n}) \xrightarrow{\text{p}} (\Rightarrow) 1.5$, и $F_{1.5}(x)$ непрерывна в точке 2.25.
Для второго критерия

$$\alpha_1(\delta_2) = \mathsf{P}_{H_1}(\overline{X} < 2.25) = \mathsf{P}_{H_1}(\sqrt{n \cdot 16}(\overline{X} - 2.25) < 0) \to \Phi_{0,1}(0) = 1/2.$$

$$\alpha_2(\delta_2) = \mathsf{P}_{H_2}(\overline{X} \geqslant 2.25) = \!\!\longrightarrow \mathsf{P}(1.5 \geqslant 2.25) = 0.$$

Снова переход к пределу объясняется слабой сходимостью в силу ЗБЧ и непрерывностью предельной ф.р. $F_{1.5}(x)$ в точке 2.25. А почему при вычислении предела $\alpha_1(\delta_2)$ мы воспользовались иными аргументами?

5. Дана выборка $X_1, \ldots, X_n$, элементы которой принимают целые неотрицательные значения $k = 0, 1, 2, 3$ и т. д. Рассматриваются две простые гипотезы: $H_1:\ \mathsf{P}\{X_i{=}k\} = \dfrac{1}{k!}\, e^{-1}$, $\quad H_2:\ \mathsf{P}\{X_i{=}k\} = \dfrac{3^k}{k!}\, e^{-3}$.

Критерий δ предписывает принимать гипотезу H_1, если $\max(X_1, \ldots, X_n) < 2$, и альтернативу H_2 — в противном случае. Найти минимальный размер выборки, при котором мощность этого критерия превышает заданное значение β.

Решение. Мощность критерия равна

$$1 - \mathsf{P}_{H_2}(\delta = H_1) = 1 - \mathsf{P}_{\Pi_3}(\max(X_1, \ldots, X_n) < 2) = 1 - (\mathsf{P}_{\Pi_3}(X_1 \leqslant 1))^n = 1 - \left(4\,e^{-3}\right)^n.$$

Требуется найти наименьшее n такое, что $1 - \left(4\,e^{-3}\right)^n > \beta$. Не забудьте, что $4 < e^3$. Решая, получим $n > \dfrac{\ln(1-\beta)}{\ln 4 - 3}$. Искомое n равно целой части этого числа плюс один.