Практическая работа 4. Генерация распределений. Проверка определений известных распределений

Юрченков Иван Александрович, ассистент кафедры ПМ

2022-10-07

Постановка задачи

1. **Сгенерировать выборку нормального распределения** $Y \sim N(\mu, \sigma^2)$ **используя определение центральной предельной теоремы**.

- На основе $n \approx 10 \div 20$ равномерно распределенных случайных реализаций случайных величин образовать новую выборку по определению центральной предельной теоремы.

Если $Y_i \sim U(a_i, b_i),\ i = 1, 2, \dots, n$, где Y_i – равномерно распределенная реализация случайной величины со случайными параметрами $a_i \in \mathbb{R},\ b_i \in \mathbb{R}$, то ожидаемая нормально распределенная величина Y будет найдена как:

$$Y = \sum_{i=1}^{n} Y_i,\ i = 1, 2, \dots, n$$

- Для получившейся выборки построить гистограмму, визуализировать на гистограмме теоретическую плотность нормального распределения по несмещенным точечным оценкам $\hat{\mu}, \hat{\sigma}$.
- Провести тест на нормальное распределение с помощью критерия χ^2-Пирсона. Степени свободы рассчитывать как $k = n - 1$.

Для генерации выборок рекомендуется пользоваться встроенными в компьютерные статистические пакеты функциями генерации **равномерно распределённых случайных величин**, которые задаются с помощью параметров границ интервала генерации чисел a и b.

2. **Сгенерировать выборку** χ^2**-распределения** $R \sim \chi^2_k$ **используя определение распределения** χ^2.

- На основе Z-оценок нормально распределенных случайных реализаций случайных величин $L_i \sim N(\mu_i, \sigma_i^2)$ образовать новую выборку по определению χ^2-распределения:

$$R = \sum_{i=1}^{n} Z[L_i]^2,\ Z[L_i] = \frac{L - E[L]}{\sigma[L]},\ L_i \sim N(\mu_i, \sigma_i^2),\ i = 1, 2, \dots, n$$

- Для получившейся выборки построить гистограмму, визуализировать на гистограмме теоретическую плотность χ^2_k распределения с $k = n - 1$ степенями свободы.
- Провести тест на χ^2 с помощью критерия χ^2-Пирсона.

Для генерации **нормально распределенных реализаций** случайных величин рекомендуется пользоваться встроенными в статистические пакеты функциями для генерации значений выборки из нормального распределения, которые задаются с помощью параметров математического ожидания μ и стандатрного отклонения σ^2.

3. **Сгенерировать выборку распределения Фишера на основе определения**.

- На основе двух случайных реализаций Y_1, Y_2 случайных величин, распределенных по χ^2-распределению со степенями свободы $d1, d2$ соответственно, сгенерировать выборку, распределенную по распределению Фишера $S \sim F(d1, d2)$ в соответствии с определением:

$$S = \frac{Y_1/d_1}{Y_2/d_2},\ S \sim F(d_1, d_2).$$

- Для получившейся выборки построить гистограмму, визуализировать на гистограмме теоретическую плотность $F(d1, d2)$ распределения.
- Провести тест на распределение Фишера с помощью критерия χ^2-Пирсона.

Для генерации выборки фиксированного размера из распределения χ^2 рекомендуется пользоваться встроенными в статистические пакеты функциями для генерации случайных выборок из распределения χ^2 с df степенями свободы.

4. **Сгенерировать выборку t-распределения на основе определения**.

- На основе $n \approx 2 \div 8$ случайных реализаций $Y_1, Y_2, \dots, Y_n$ случайных величин, распределенных по стандартному нормальному распределению $Y_i \sim N(0,1),\ i = 1, 2, \dots, n$, сгенерировать выборку $T \sim t(n)$, распределенную по t-распределению Стьюдента с $df = n$ степенями свободы в соответствии с определением:

$$T = \frac{Y_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2}}, \quad Y_0 \sim N(0,1).$$

- Реализовать вычисление аналитической плотности t-распределения Стьюдента с использованием бета-функции:

$$p_t(x \mid n) = \frac{1}{\sqrt{n}\,\mathrm{B}(\frac{1}{2}, \frac{n}{2})}\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}},$$

где

$$\mathrm{B}(x, y) = \int\limits_0^1 t^{x-1}(1-t)^{y-1}\, dt,$$

определённая при Re $x > 0$, Re $y > 0$.

- Для получившейся выборки построить гистограмму, визуализировать на гистограмме теоретическую плотность $t(n)$.
- Для получившейся выбрки провести тест на t-распределение Стьюдента с помощью критерия χ^2-Пирсона, используя в качестве функции вероятности распределения $P_t(x \mid n)$:

$$P_t(x \mid n) = \int_{-\infty}^{x} p_t(z \mid n) dz.$$

5. Для всех заданий количество генерируемых значений выборки установить равным $N \approx 100 \div 1000$. Уровень надежности для критерия χ^2-Пирсона или метода анаморфоз $\gamma = 0.95$.

Вопросы на защиту практической работы

1. Центральная предельная теорема. Реализации случайно распределенных величин. Независимые величины. Степени свободы суммы независимо распределенных величин.
2. Определение нормального распределения. Спрямление для координат нормального распределения. Определение параметров нормального распределения через точечные оценки. Определение параметров нормального распределения, образованного суммой независимых величин, через ЦПТ.
3. Определение распределения Фишера. Аналитические формулы математического ожидания и дисперсии расределения Фишера.
4. t-распределение Стьюдента. Аппроксимации и определение функции плотности. Смесь нормально расределенных величин. Определение Z-оценок.