Лаб-2 Щепкин Егор В-20

1.2. Будем доказывать от обратного.

Пусть $X = Y + Z$.

Очевидно, что Y и Z также дискретные величины.

А так как распределение не вырожд[ено], то они принимают хотя бы 2 различных значения.

Б.О.О. $y_1 < y_2$ и $z_1 < z_2$ соответственно

но $y_1 + z_1 < y_2 + z_1 < y_2 + z_2$ – три различных значения в сумме Y и Z, а должно быть 2 противоречие. ЧТД

2.2. $\forall$ $f_R(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \mathbb{1}\{x > 0\}$

$$f_\Theta(x) = \frac{1}{2\pi} \mathbb{1}\{x \in [0; 2\pi]\}$$

В силу независимости R и Θ

$$f_{R,\Theta}(x, y) = \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}} \frac{1}{2\pi}$$

$X = R\sin\theta$

$Y = R\cos\theta$

Воспользуемся формулой преобразования плотности

$$f_{R,\theta}(R, r, \theta) = \cancel{f_{X,Y}(r\cos\theta, r}$$

$$= f_{X,Y}(r\sin\theta, r\cos\theta)\,|\det D(r\sin\theta, r\cos\theta)|$$

$$\det D(r\sin\theta, r\cos\theta) = \begin{vmatrix} \frac{\partial r\sin\theta}{\partial r} & \frac{\partial r\sin\theta}{\partial\theta} \\ \frac{\partial r\cos\theta}{\partial r} & \frac{\partial r\cos\theta}{\partial\theta} \end{vmatrix}$$

$$= \begin{vmatrix} \sin\theta & r\cos\theta \\ \cos\theta & -r\sin\theta \end{vmatrix} = r$$

$$f_{X,Y}(r\sin\theta, r\cos\theta) = \frac{f_{R,\theta}(r, \theta)}{r} = \frac{\frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \frac{1}{2\pi}}{r} =$$

$$= \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{r^2}{2\sigma^2}}$$

$$R^2 = X^2 + Y^2 = R^2(\cos^2\theta + \sin^2\theta) \qquad \boxed{r^2 = x^2 + y^2}$$

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}} =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{y^2}{2\sigma^2}} =$$

$$f_X(x) \cdot f_Y(y) \quad \text{— ЧТД.}$$

И как несложно заметить f_x и f_y — нормальные распределение.

$$X \sim N(0,\sigma) \qquad Y \sim N(0,\sigma)$$

3.4

$$p(x) = \frac{e^{-|x|}}{2}(1-e^{-1})^{-1}\,\mathbb{1}(|x| \le 1)$$

$F(x) = ?$

$$F(x) = \begin{cases} \int_{-1}^{x} \frac{e^{+x'}}{2}(1-e^{-1})^{-1}\,dx', & -1 \le x < 0 \\ \frac{1}{2} + \int_{0}^{x} \frac{e^{-x'}}{2}(1-e^{-1})^{-1}\,dx', & 0 \le x \le 1 \end{cases}$$

Это верно в силу симметричности $p(x)$ относительно 0.

$$F(x) = \begin{cases} \frac{(1-e^{-1})^{-1}}{2}\int_{-1}^{x} e^{x'}\,dx', & -1 \le x < 0 \\ \frac{1}{2} + \frac{(1-e^{-1})^{-1}}{2}\int_{0}^{x} e^{-x'}\,dx', & 0 \le x \le 1 \end{cases}$$

$$F(x) = \begin{cases} \frac{(1-e^{-1})^{-1}}{2}\left(e^{x} - \frac{1}{e}\right), & -1 \le x < 0 \\ \frac{1}{2} + \frac{(1-e^{-1})^{-1}}{2}\left(1 - e^{-x}\right) & 0 \le x \le 1 \end{cases}$$

$F^{-1}(x) =$

Если $x < \frac{1}{2}$, то

$$x = \frac{(1-e^{-1})^{-1}}{2}\left(e^{x'} - \frac{1}{e}\right)$$

$$x \cdot 2 \cdot (1-e^{-1}) + \frac{1}{e} = e^{x'}$$

$$x' = \ln\left(x \cdot 2 \cdot (1-e^{-1}) + \frac{1}{e}\right)$$

Если $x \geq \frac{1}{2}$, то

$$x = \frac{1}{2} + \frac{(1-e^{-1})^{-1}}{2}\left(1 - e^{-x'}\right)$$

$$\left(x - \frac{1}{2}\right) 2 \cdot (1-e^{-1}) = 1 - e^{-x'}$$

$$1 - \left(x - \frac{1}{2}\right) \cdot 2 (1-e^{-1}) = e^{-x'}$$

$$-\ln\left(1 - \left(x - \frac{1}{2}\right) \cdot 2 (1-e^{-1})\right) = x'$$

$$F^{-1}(x) = \begin{cases} \ln\left(x \cdot 2 \cdot (1-e^{-1}) + \frac{1}{e}\right), & x \leq \frac{1}{2} \\ -\ln\left(1 - \left(x - \frac{1}{2}\right) \cdot 2 \cdot (1-e^{-1})\right) \end{cases}$$

Для 3 теста я взял Ratio of Uniform
т.к. он уже реализован в SciPy
Эксперименты показали, что
генерация 10000 чисел через rv continues
заняло 30 секунд
через F^{-1}: ~~мс~~ ≈ 100 миллисекунд
через RoF: 10 секунд

Вывод: наиболее быстрый вариант использовать F^{-1}, но вычисления F^{-1} могут оказаться сложными, в данном случае нам повезло найти явный вид

4.5 $Exp(\lambda)$, $\lambda \in [1; 5]$

$X_1, X_2, \dots, X_n$ - i.i.d. $\quad X_i = Exp(\lambda)$ для фикс. λ.

$$\overline{X}_n = \frac{\sum X_i}{n}$$

μ_λ - мат ожидание

Найти $n_{\delta,\varepsilon}$: $\forall \lambda \in [1, 5]$ и $\delta, \varepsilon \in (0, 1)$

$$P\left(\left|\overline{X}_{n_{\delta,\varepsilon}} - \mu_\lambda\right| \leq \varepsilon\right) \geq 1 - \delta$$

$$E\overline{X}_{n_{\delta,\varepsilon}} = \frac{\sum_{i=1}^{n} EX_i}{n} = EX$$

$$Var\, \overline{X}_{n_{\delta,\varepsilon}} = \frac{Var \sum^n X_i}{n^2} = \frac{\sum Var X_i}{n^2} = \frac{n\, Var X}{n^2} = \frac{Var X}{n}$$

$$E\overline{X}_{n_{\delta,\varepsilon}} = EX_\lambda = \frac{1}{\lambda} = \mu_\lambda$$

$$Var\, \overline{X}_{n_{\delta,\varepsilon}} = \frac{Var X}{n} = \frac{1}{\lambda^2 n}$$

Неравенство Чебышёва.

$$P\left(\left|\overline{X_{n_{\delta,\varepsilon}}} - \mu_\lambda\right| > \varepsilon\right) \le \frac{\operatorname{Var} \overline{X_{n_{\delta,\varepsilon}}}}{\varepsilon^2}$$

$$\| \; P(\ldots)$$

$$1 - P\left(\left|\overline{X_{n_{\delta,\varepsilon}}} - \mu_\lambda\right| \le \varepsilon\right) = 1 - P(\ldots)$$

$$P\left(\left|\overline{X_{n_{\delta,\varepsilon}}} - \mu_\lambda\right| \le \varepsilon\right) \ge 1 - \frac{\operatorname{Var} \overline{X_{n_{\delta,\varepsilon}}}}{\varepsilon^2} =$$

$$= 1 - \frac{\operatorname{Var} X}{n_{\delta,\varepsilon}\, \varepsilon^2} \ge 1 - \delta$$

$$\frac{\operatorname{Var} X}{n_{\delta,\varepsilon}\, \varepsilon^2} \le \delta$$

$$n_{\delta,\varepsilon} \ge \frac{\operatorname{Var} X}{\delta, \varepsilon^2} = \frac{1}{\lambda^2 \varepsilon^2 \delta}$$

ЦПТ.

$$P\left(\left|\overline{X_{n_{\delta,\varepsilon}}} - \mu_\lambda\right| \le \varepsilon\right) = P\left(\left|\frac{S_{n_{\delta,\varepsilon}}}{n_{\delta,\varepsilon}} - \mu_\lambda\right| \le \varepsilon\right) =$$

$$= P\left(\left|\frac{S_{n_{\delta,\varepsilon}} - \mu_\lambda n_{\delta,\varepsilon}}{\sqrt{n_{\delta,\varepsilon}}\, \sigma}\right| \le \frac{\sqrt{n_{\delta,\varepsilon}}\, \varepsilon}{\sigma}\right) \approx$$

$$\Phi\left(\frac{\sqrt{n_{\delta,\varepsilon}}\, \varepsilon}{\delta}\right) \ge 1 - \delta \qquad \Phi\left(\lambda \sqrt{n_{\delta,\varepsilon}}\, \varepsilon\right) \ge 1 - \delta$$

Можно найти бинарным поиском

Так как Φ возрастает.

После тестирования Чебышёв даёт 100% попадание в нужный интервал, а ЦПТ даёт $\delta \approx 13\%$, что в 2,5 раза больше ожидаемого. Это связано с тем, что ~~при~~ n выходит маленьким в районе 1000 и у ЦПТ здесь большая погрешность. Но значения n в десятки раз или даже сотни раз меньше чем даёт Чебышёв.

Я дополнительно проверил ЦПТ на $n \cdot 2$ и теперь $\delta \approx 5\%$ как и нужно.

Вывод: Чебышёв даёт гарантию выполнения условия, а ЦПТ менее точные но меньшие значения.