Лекция 3

Выборка из нормального распределения и теорема Фишера

График «квантиль-квантиль»

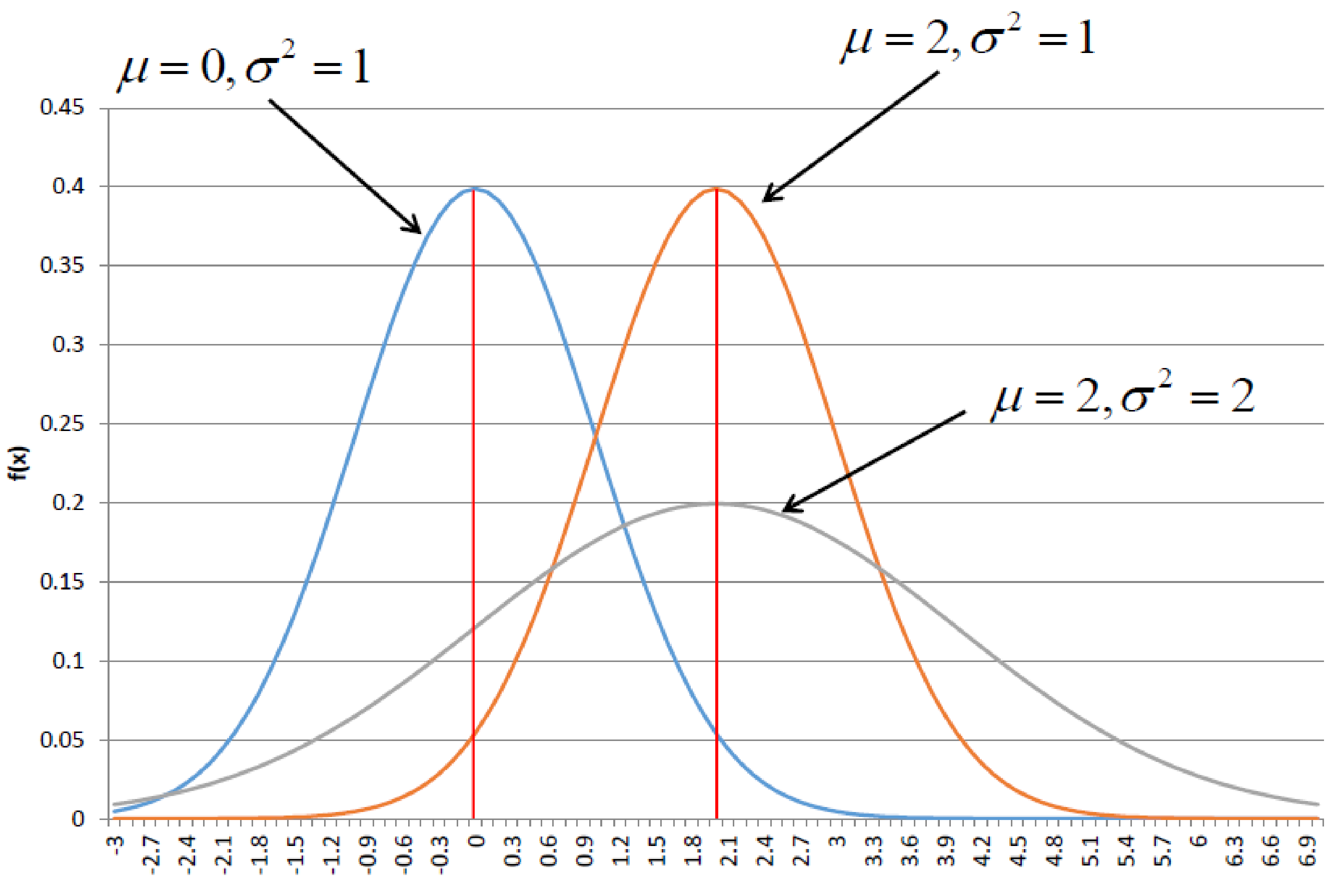
Распределение выборочного среднего и выборочной доли

Напоминалка 1: нормальное распределение

Случайная величина X имеет нормальное распределение с параметрами μ и $\sigma^2>0$, если её функция плотности имеет вид

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Обозначение: $X\sim N(\mu,\sigma^2)$.



Напоминалка 1: нормальное распределение

Свойства нормальных случайных величин:

1°. Пусть $X \sim N(\mu, \sigma^2)$.

Тогда $\mathrm{E}(X) = \mu$, $\mathrm{D}(X) = \sigma^2$.

2°. Пусть $X \sim N(\mu, \sigma^2)$, $\quad Y = aX + b$, $a \neq 0$.

Тогда $Y \sim N(a\mu + b, a^2\sigma^2)$.

3°. Пусть X, Y независимы, $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$.

Тогда $X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

Важный частный случай свойства 2:

$$Z = \frac{X - \mu}{\sigma} \sim N(0,\, 1).$$

стандартное нормальное
распределение

Напоминалка 2: распределение хи-квадрат

Пусть случайные величины $Z_1, \ldots, Z_k$ независимы, $Z_i \sim N(0, 1)$.

Распределение случайной величины

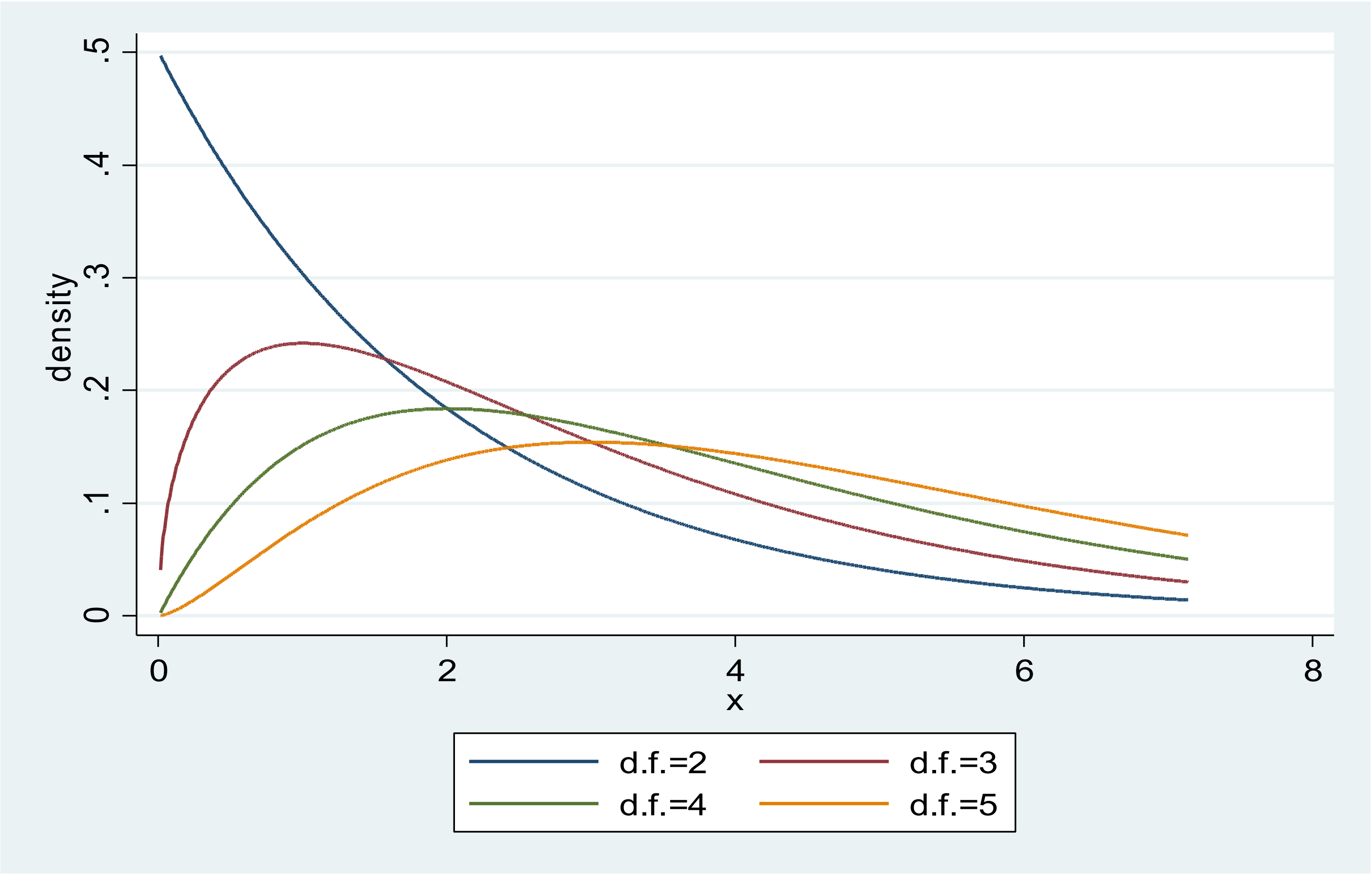
$$Y = Z_1^2 + \ldots + Z_k^2$$

Называется распределением хи-квадрат с k степенями свободы.

Обозначение: $Y \sim \chi_k^2$.

$\mathrm{E}(Y) = k, \ \mathrm{D}(Y) = 2k.$

плотность для 2, 3, 4, 5 степеней свободы:



Напоминалка 2: распределение хи-квадрат

Пусть случайные величины $Z_1, \ldots, Z_k$ независимы, $Z_i \sim N(0, 1)$.

Распределение случайной величины

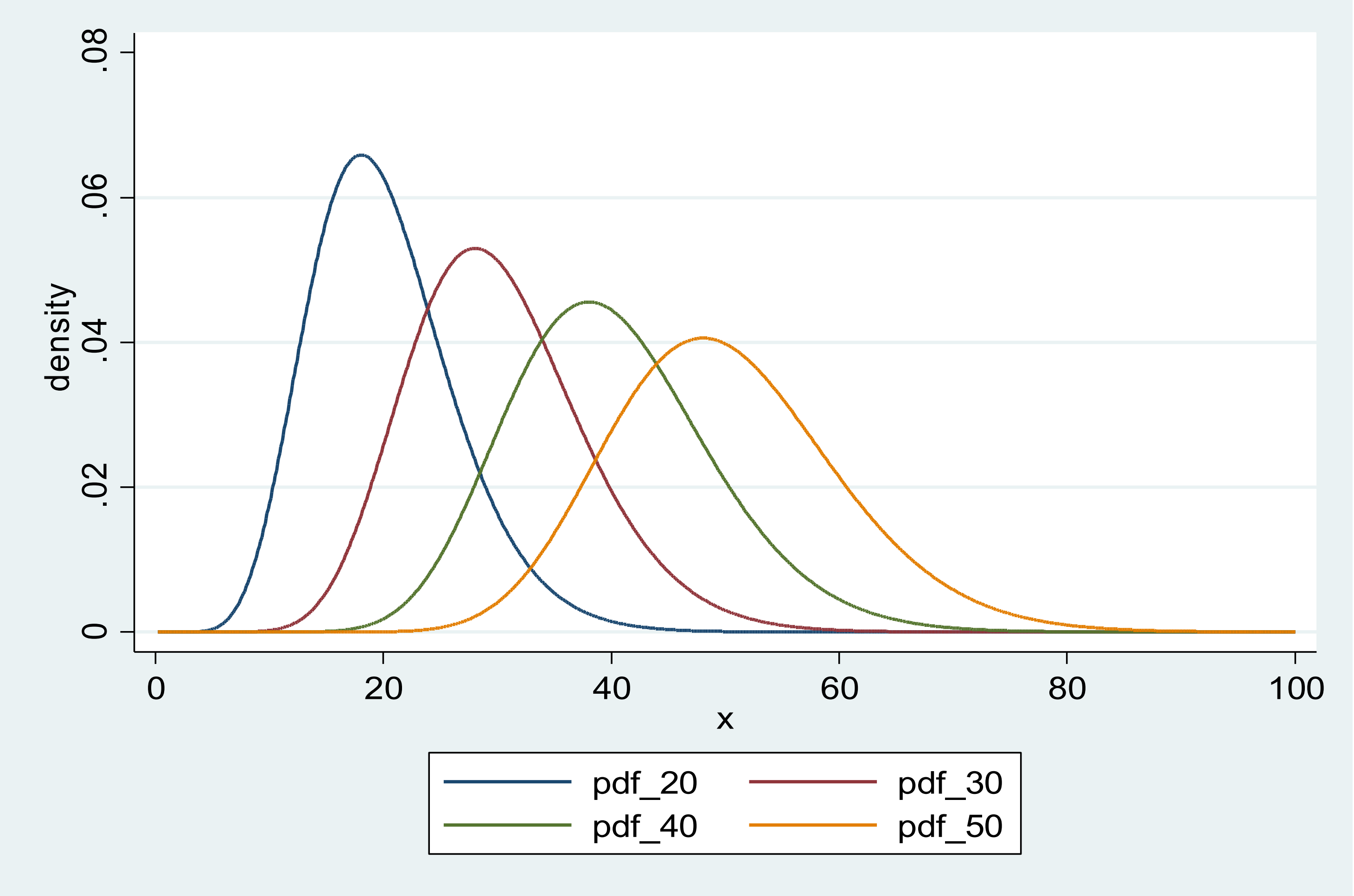
$$Y = Z_1^2 + \ldots + Z_k^2$$

Называется распределением хи-квадрат с k степенями свободы.

Обозначение: $Y \sim \chi_k^2$.

$$\mathrm{E}(Y) = k, \quad \mathrm{D}(Y) = 2k.$$

а теперь 20, 30, 40, 50 степеней свободы:



Выборка из нормального распределения. Теорема Фишера.

Пусть $X_1, \ldots, X_n$ независимы, $X_i \sim N(\mu, \sigma^2)$.

Тогда:

(1) $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$;

(2) $\dfrac{nS^2}{\sigma^2} = \dfrac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1}$;

(3) $\bar{X}$ и S^2 независимы. ($\bar{X}$ и $\hat{\sigma}^2$ тоже независимы)

Напоминалка:

$$\bar{X} = \frac{X_1 + \ldots + X_n}{n}; \qquad S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2; \qquad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

какой толк?

Теорема Фишера, пример

В упаковке должно содержаться в среднем 100 грамм чая со стандартным отклонением не более 3 грамм. Время от времени отдел контроля качества отбирает 16 упаковок для проверки и рассчитывает среднее $\bar{X}$ и оценку стандартного отклонения $\hat{\sigma}$.

Партия чая проходит контроль качества, если выполняются два условия:

(1) $98 \leq \bar{X} \leq 102;$

(2) $\hat{\sigma} \leq 4.28.$

Если хотя бы одно из них нарушается, процесс упаковки останавливается для переналадки.

Предположим, что масса чая распределена нормально со средним 100 грамм и стандартным отклонением 3 грамма. С какой вероятностью процесс упаковки будет остановлен (произойдёт ложная тревога)?

Теорема Фишера, пример

В упаковке должно содержаться в среднем 100 грамм чая со стандартным отклонением не более 3 грамм. Время от времени отдел контроля качества отбирает 16 упаковок для проверки и рассчитывает среднее $\bar{X}$ и оценку стандартного отклонения $\hat{\sigma}$.

Партия чая проходит контроль качества, если выполняются два условия:

(1) $98 \le \bar{X} \le 102$;

(2) $\hat{\sigma} \le 4.28$.

Если хотя бы одно из них нарушается, процесс упаковки останавливается для переналадки.

Предположим, что масса чая распределена нормально со средним 100 грамм и стандартным отклонением 3 грамма. С какой вероятностью процесс упаковки будет остановлен (произойдёт ложная тревога)?

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \ldots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right).$

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \ldots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Сначала разберёмся со средним.

$$\bar{X} \sim N\left(100,\ \frac{9}{16}\right)$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Сначала разберёмся со средним.

$$\bar{X} \sim N\left(100,\ \frac{9}{16}\right)$$

Центрируем и нормируем:

$$Z = \frac{\bar{X} - 100}{\sqrt{9/16}} = \frac{\bar{X} - 100}{3/4} \sim N(0,1).$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \ldots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right).$

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Сначала разберёмся со средним.

$$\bar{X} \sim N\left(100,\ \frac{9}{16}\right)$$

Центрируем и нормируем:

$$Z = \frac{\bar{X} - 100}{\sqrt{9/16}} = \frac{\bar{X} - 100}{3/4} \sim N(0,1).$$

$$P\left(98 \le \bar{X} \le 102\right) = P\left(\frac{98 - 100}{3/4} \le Z \le \frac{102 - 100}{3/4}\right) =$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Сначала разберёмся со средним.

$$\bar{X} \sim N\left(100,\ \frac{9}{16}\right)$$

Центрируем и нормируем:

$$Z = \frac{\bar{X} - 100}{\sqrt{9/16}} = \frac{\bar{X} - 100}{3/4} \sim N(0,1).$$

$$P\left(98 \le \bar{X} \le 102\right) = P\left(\frac{98-100}{3/4} \le Z \le \frac{102-100}{3/4}\right) = P\left(-\frac{8}{3} \le Z \le \frac{8}{3}\right) =$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \ldots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1-P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Сначала разберёмся со средним.

$$\bar{X} \sim N\left(100, \frac{9}{16}\right)$$

Центрируем и нормируем:

$$Z = \frac{\bar{X}-100}{\sqrt{9/16}} = \frac{\bar{X}-100}{3/4} \sim N(0,1).$$

волшебная функция

$$P\left(98 \le \bar{X} \le 102\right) = P\left(\frac{98-100}{3/4} \le Z \le \frac{102-100}{3/4}\right) = P\left(-\frac{8}{3} \le Z \le \frac{8}{3}\right) = 0.9923.$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Теперь займёмся стандартным отклонением.

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} = \frac{15\hat{\sigma}^2}{9} \sim \chi^2_{15}.$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Теперь займёмся стандартным отклонением.

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} = \frac{15\,\hat{\sigma}^2}{9} \sim \chi^2_{15}.$$

$$P\left(\hat{\sigma} \le 4.28\right) = P\left(\hat{\sigma}^2 \le 4.28^2\right) =$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Теперь займёмся стандартным отклонением.

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} = \frac{15\,\hat{\sigma}^2}{9} \sim \chi^2_{15}.$$

$$P\left(\hat{\sigma} \le 4.28\right) = P\left(\hat{\sigma}^2 \le 4.28^2\right) = P\left(\underbrace{\frac{15\,\hat{\sigma}^2}{9}}_{\chi^2_{15}} \le \frac{15 \times 4.28^2}{9}\right) =$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \dots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Теперь займёмся стандартным отклонением.

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} = \frac{15\hat{\sigma}^2}{9} \sim \chi^2_{15}.$$

$$P\left(\hat{\sigma} \le 4.28\right) = P\left(\hat{\sigma}^2 \le 4.28^2\right) = P\left(\underbrace{\frac{15\hat{\sigma}^2}{9}}_{\chi^2_{15}} \le \frac{15 \times 4.28^2}{9}\right) = P\left(\frac{15\hat{\sigma}^2}{9} \le 30.53\right) =$$

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \ldots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Теперь займёмся стандартным отклонением.

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} = \frac{15\hat{\sigma}^2}{9} \sim \chi^2_{15}.$$

$$P\left(\hat{\sigma} \le 4.28\right) = P\left(\hat{\sigma}^2 \le 4.28^2\right) = P\left(\underbrace{\frac{15\hat{\sigma}^2}{9}}_{\chi^2_{15}} \le \frac{15 \times 4.28^2}{9}\right) = P\left(\frac{15\hat{\sigma}^2}{9} \le 30.53\right) = 0.9899.$$

снова волшебная функция

Теорема Фишера, пример

Решение. Пусть X_i — вес упаковки i в выборке. Тогда $X_1, \ldots, X_{16}$ независимы,

$$X_i \sim N\left(100,\ 3^2\right).$$

Нужно найти $1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right)$.

По т. Фишера $\bar{X}$ и $\hat{\sigma}^2$ независимы, так что

$$P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = P\left(98 \le \bar{X} \le 102\right) P\left(\hat{\sigma} \le 4.28\right).$$

Итак,

$$P\left(98 \le \bar{X} \le 102\right) = 0.9923.$$

$$P\left(\hat{\sigma} \le 4.28\right) = 0.9899.$$

Значит,

$$1 - P\left(\{98 \le \bar{X} \le 102\} \cap \{\hat{\sigma} \le 4.28\}\right) = 1 - 0.9923 \times 0.9899 = 0.0177.$$

Ответ. Вероятность ложной тревоги составляет 1.77%

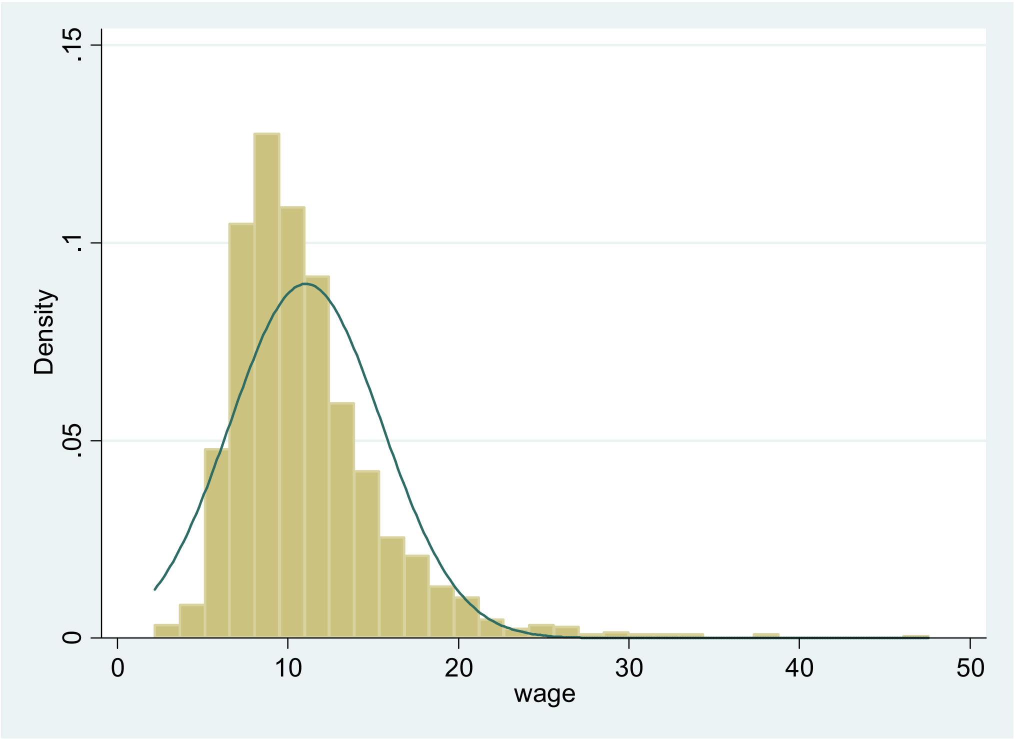
Как понять, что выборка взята именно из нормальной генеральной совокупности?

Гистограмма

Можно построить гистограмму и сравнить её с нормальной функцией плотности.

Пример. Данные о зарплатах 1492 рабочих в Бельгии, 1994 год.

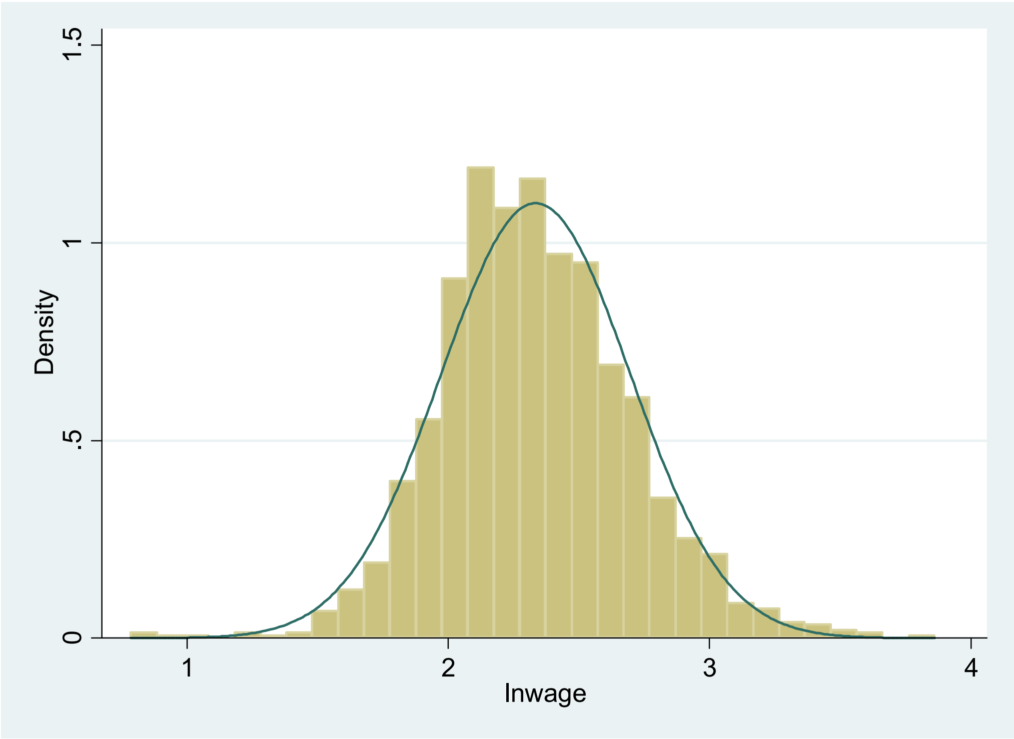
исходные данные



$$\bar{X} = 11.05$$

$$\hat{\sigma}_X^2 = 19.81$$

логарифм зарплаты



$$\bar{X}^* = 2.33$$

$$\hat{\sigma}_X^2* = 0.13$$

Недостатки:
- нужно много наблюдений,
- некоторые существенные отклонения от нормальности сложно заметить.

График «квантиль-квантиль»

(Q-Q plot)

1. Оцениваем параметры нормального распределения по выборке:

$$\bar{X} = \frac{X_1 + \ldots + X_n}{n}; \qquad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

(есть варианты)

2. Рассчитываем выборочные квантили порядка $\frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n}{n+1}$

- это будет просто упорядоченный по возрастанию ряд $X_{(1)}, \ldots, X_{(n)}$.

наименьшее наибольшее

$$\hat{Q}\left(\frac{1}{n+1}\right) = X_{(1)},\ \hat{Q}\left(\frac{2}{n+1}\right) = X_{(2)},\ \ldots$$

3. Рассчитываем квантили нормального распределения с параметрами $\mu = \bar{X}$ и $\sigma^2 = \hat{\sigma}^2$

(«теоретические» квантили):

$$Q\left(\frac{1}{n+1}\right) = \bar{X} + \hat{\sigma}\,\Phi^{-1}\left(\frac{1}{n+1}\right),\ Q\left(\frac{2}{n+1}\right) = \bar{X} + \hat{\sigma}\,\Phi^{-1}\left(\frac{2}{n+1}\right), \ldots$$

4. Строим график в осях $\left(Q, \hat{Q}\right)$ - для выборок из нормального распределения должно

быть $Q(p) \approx \hat{Q}(p)$ (график выстраивается вдоль биссектрисы угла, образованного осями).

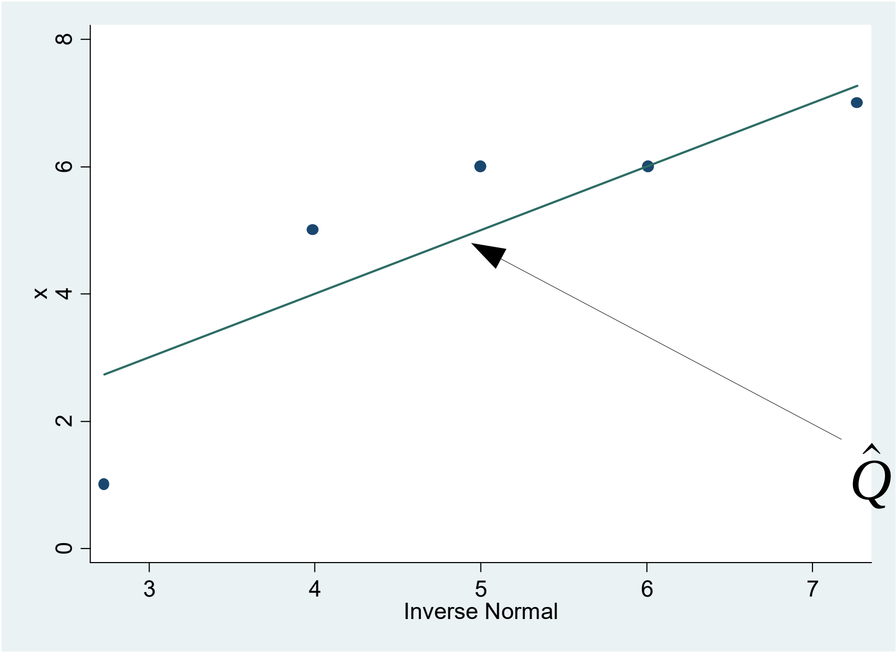
Пример

Возьмём выборку 5 1 6 6 7

Оценим среднее и дисперсию: $\bar{X}=5,\ \hat{\sigma}^2=5.5 \quad \Rightarrow \quad \hat{\sigma}=\sqrt{5.5}\approx 2.35.$

порядок квантили, p	выборочная квантиль $\hat{Q}(p)$	квантиль N(0, 1) $\Phi^{-1}(p)$	теоретическая квантиль $Q(p)=5+2.35\,\Phi^{-1}(p)$
$^1/_6$	1	-0.97	2.73
$^2/_6$	5	-0.43	3.99
$^3/_6$	6	0.00	5.00
$^4/_6$	6	0.43	6.01
$^5/_6$	7	0.97	7.27

График «квантиль-квантиль»:

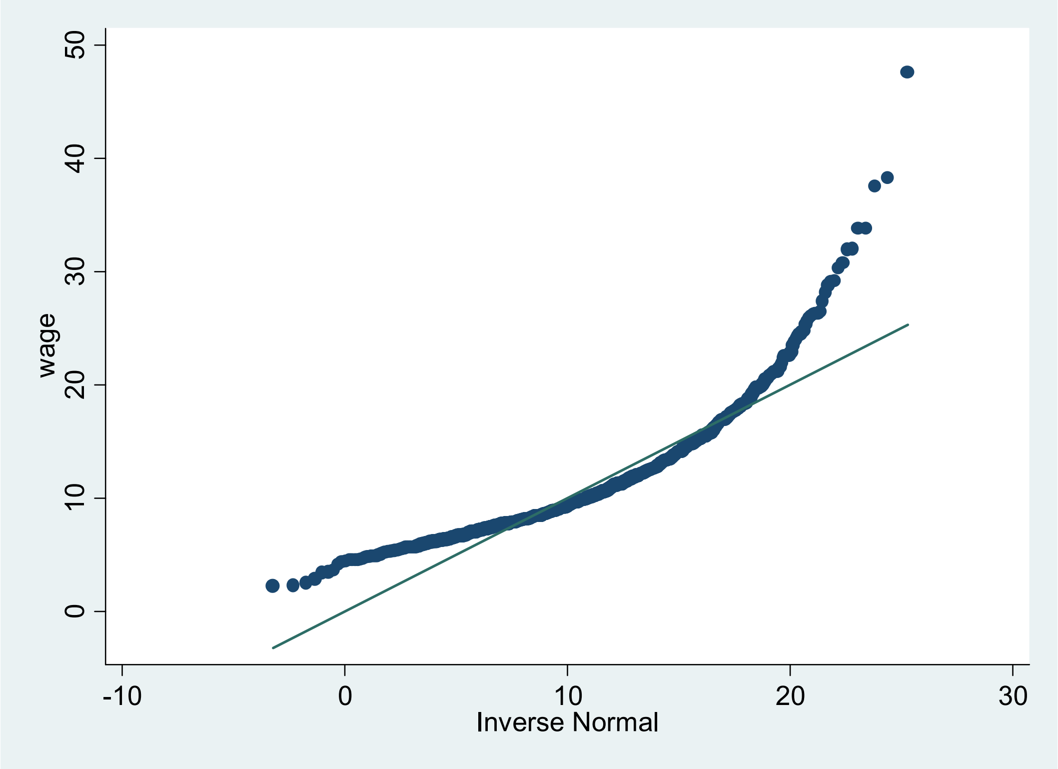


$\hat{Q}(p)=Q(p)$

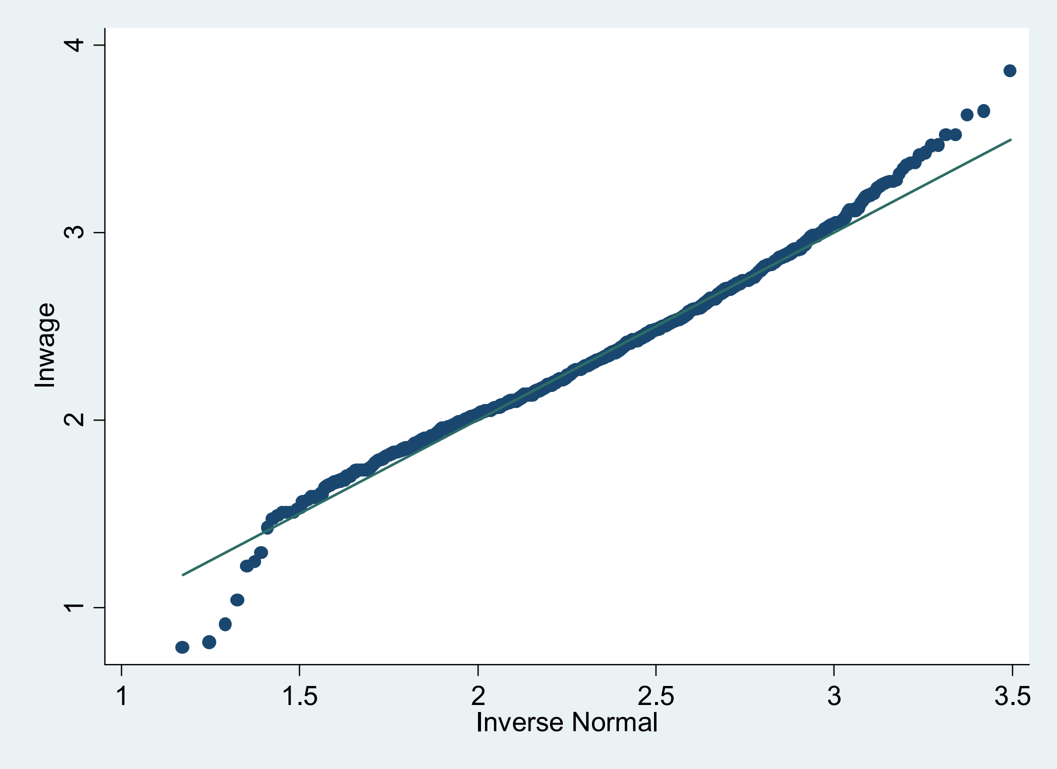
По горизонтали — теоретические квантили, по вертикали — выборочные.

График «квантиль-квантиль» для зарплат в Бельгии

исходные данные



логарифмы



Среднее в выборке из не нормальной генеральной совокупности

Есть случайная выборка $X_1, \ldots, X_n$, такая что

$$X_i \sim \text{i.i.d.}; \quad \mathrm{E}\left(X_i\right) = \mu; \quad \mathrm{D}\left(X_i\right) = \sigma^2, \; 0 < \sigma^2 < \infty.$$

Тогда:

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \overset{\text{asy}}{\sim} N(0,1).$$

по центральной предельной теореме

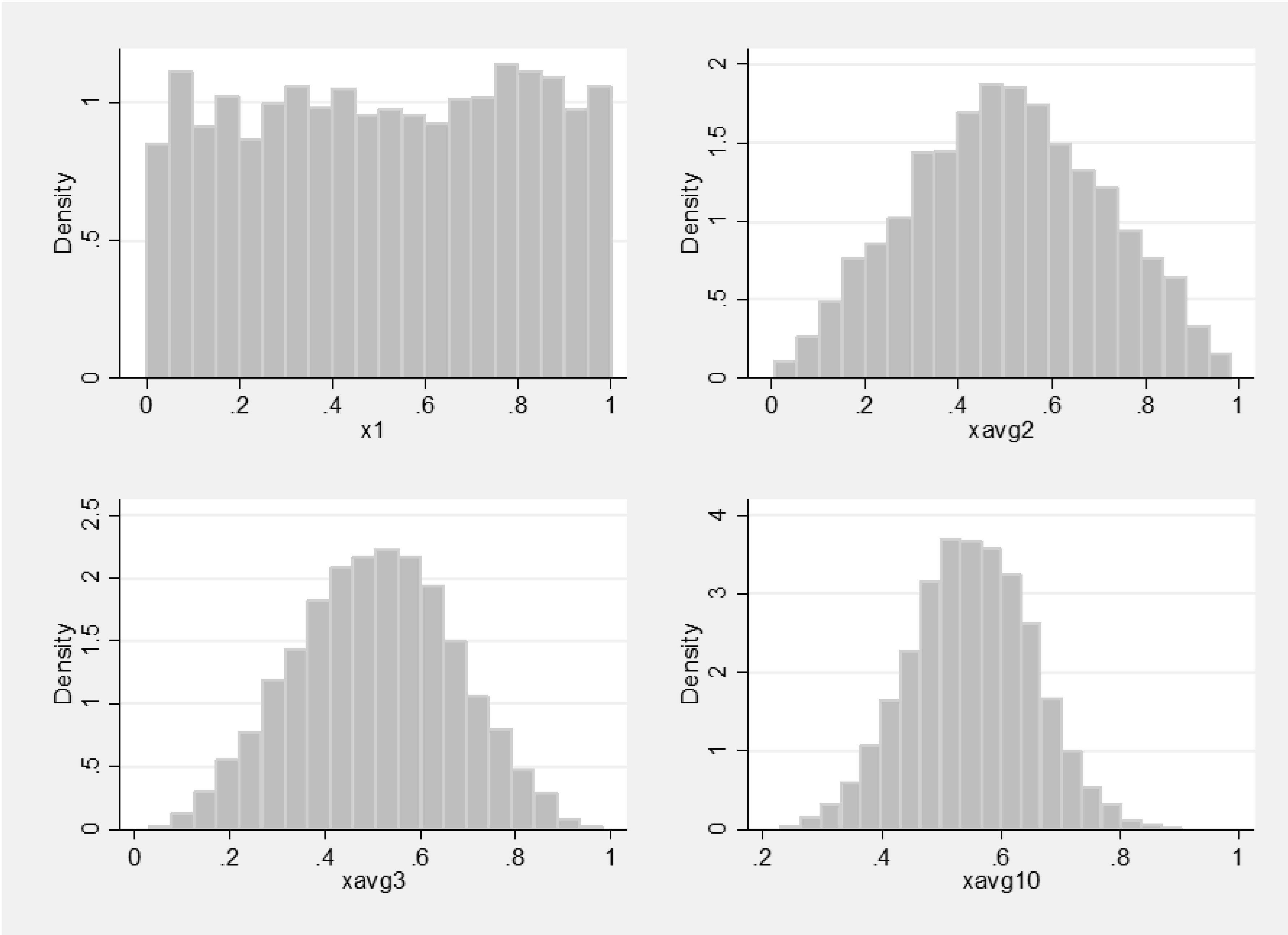
Мораль. На больших выборках можно пользоваться нормальным приближением:

$$\boxed{\bar{X} \overset{\text{app}}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right).}$$

asy — asymptotically (асимптотически)
app — approximately (приближённо)

Пример: распределение среднего в выборке из равномерного распределения

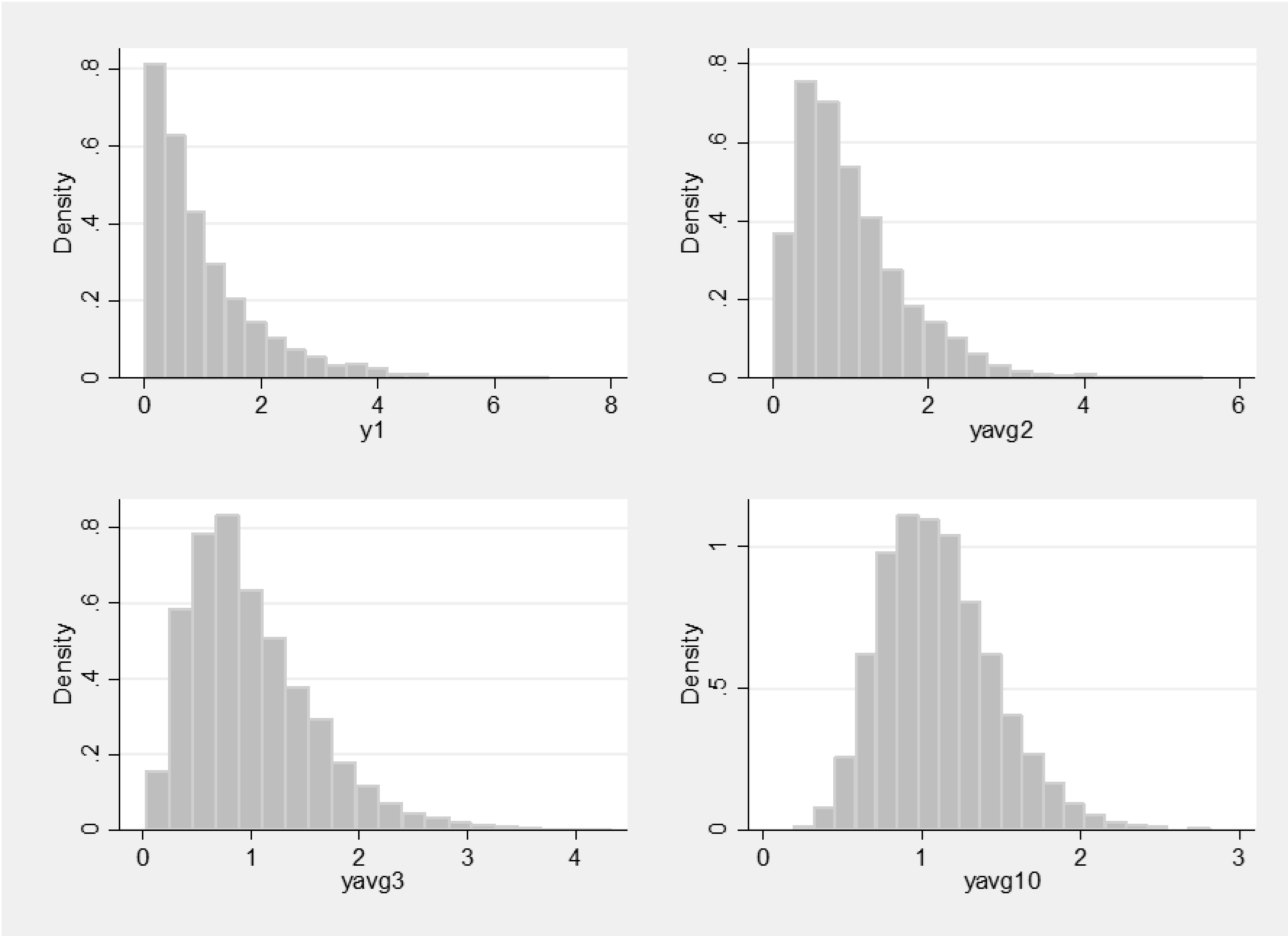
$n = 1$ $n = 2$



$n = 3$ $n = 10$

Пример: распределение среднего в выборке из показательного распределения

$n = 1$ $n = 2$



$n = 3$ $n = 10$

Распределение выборочной доли

Пусть $X_1, \ldots, X_n$ независимы,

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

Вспомним характеристики:

$$\mathrm{E}(X_i) = p; \quad \mathrm{D}(X_i) = p(1-p).$$

Выборочная доля:

$$\hat{p} = \frac{X_1 + \ldots + X_n}{n}.$$

Распределение выборочной доли:

$$\boxed{\begin{gathered} \frac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \overset{\text{asy}}{\sim} N(0,1). \\[2em] \hat{p} \overset{\text{app}}{\sim} N\left(p, \frac{p(1-p)}{n}\right). \end{gathered}}$$

Пример

Перед выборами проводится опрос, цель которого — определение доли избирателей, поддерживающих кандидата А. Планируется опросить 100 человек.

Какова вероятность того, что большинство опрошенных выскажутся в поддержку кандидата А, если его поддерживают 40% всех избирателей?

Пример

Перед выборами проводится опрос, цель которого — определение доли избирателей, поддерживающих кандидата А. Планируется опросить 100 человек.

Какова вероятность того, что большинство опрошенных выскажутся в поддержку кандидата А, если его поддерживают 40% всех избирателей?

Решение. По условию $n = 100$, $p = 0.4$, так что

$$\hat{p} \overset{\text{app}}{\sim} N\left(0.4, \frac{0.4(1-0.4)}{100}\right).$$

Пример

Перед выборами проводится опрос, цель которого — определение доли избирателей, поддерживающих кандидата А. Планируется опросить 100 человек.

Какова вероятность того, что большинство опрошенных выскажутся в поддержку кандидата А, если его поддерживают 40% всех избирателей?

Решение. По условию $n = 100$, $p = 0.4$, так что

$$\hat{p} \overset{\text{app}}{\sim} N\left(0.4, \frac{0.4(1-0.4)}{100}\right).$$

$$P(\hat{p} > 0.5) = P\left(\overbrace{\frac{\hat{p}-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}}}^{N(0,\,1)} > \frac{0.5-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}}\right)$$

Пример

Перед выборами проводится опрос, цель которого — определение доли избирателей, поддерживающих кандидата А. Планируется опросить 100 человек.

Какова вероятность того, что большинство опрошенных выскажутся в поддержку кандидата А, если его поддерживают 40% всех избирателей?

Решение. По условию $n = 100$, $p = 0.4$, так что

$$\hat{p} \overset{\text{app}}{\sim} N\left(0.4, \frac{0.4(1-0.4)}{100}\right).$$

$$P(\hat{p} > 0.5) = P\left(\overbrace{\frac{\hat{p}-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}}}^{N(0,\,1)} > \frac{0.5-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}}\right) = P\left(\frac{\hat{p}-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}} > 2.04\right)$$

Пример

Перед выборами проводится опрос, цель которого — определение доли избирателей, поддерживающих кандидата А. Планируется опросить 100 человек.

Какова вероятность того, что большинство опрошенных выскажутся в поддержку кандидата А, если его поддерживают 40% всех избирателей?

Решение. По условию $n = 100$, $p = 0.4$, так что

$$\hat{p} \overset{\text{app}}{\sim} N\left(0.4, \frac{0.4(1-0.4)}{100}\right).$$

$$P(\hat{p} > 0.5) = P\left(\overbrace{\frac{\hat{p}-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}}}^{N(0,\,1)} > \frac{0.5-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}}\right) = P\left(\frac{\hat{p}-0.4}{\sqrt{\frac{0.4(1-0.4)}{100}}} > 2.04\right) \approx 0.0206.$$

Ответ. 0.0206.

Следующая лекция

Доверительные интервалы для параметров нормальной генеральной совокупности