Теория Вероятностей и Статистика

Доверительные интервалы

Потанин Богдан Станиславович

старший преподаватель, научный сотрудник, кандидат экономических наук

2022-2023

- Ранее мы оценивали параметры распределения и получали так называемые **точечные оценки**.

- Ранее мы оценивали параметры распределения и получали так называемые **точечные оценки**.
- Однако, на практике, нас зачастую может интересовать не только истинное значение параметра, но и некоторый диапазон значений, которому соответствующий параметр может принадлежать с высокой вероятность. Процедуру нахождения такого диапазона часто именуют **интервальным оцениванием**.

- Ранее мы оценивали параметры распределения и получали так называемые **точечные оценки**.
- Однако, на практике, нас зачастую может интересовать не только истинное значение параметра, но и некоторый диапазон значений, которому соответствующий параметр может принадлежать с высокой вероятность. Процедуру нахождения такого диапазона часто именуют **интервальным оцениванием**.
- Например, нам может быть интересно оценить не только средний доход по стране, но и диапазон значений, в который средний доход попадает с вероятностью 0.99.

- Рассмотрим выборку $X = (X_1, ..., X_n)$ из распределения с параметром θ.

- Рассмотрим выборку $X = (X_1, ..., X_n)$ из распределения с параметром θ.
- **Доверительным интервалом** с уровнем доверия $(1-\gamma)$, где $\gamma \in (0,1)$, называется интервал $[T_1(X), T_2(X)]$, такой, что при любом допустимом значении параметра θ:

$$P\left(\theta \in \left[T_1\left(X\right), T_2\left(X\right)\right)\right] = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

- Рассмотрим выборку $X = (X_1, ..., X_n)$ из распределения с параметром θ.
- **Доверительным интервалом** с уровнем доверия $(1-\gamma)$, где $\gamma \in (0,1)$, называется интервал $[T_1(X), T_2(X)]$, такой, что при любом допустимом значении параметра θ:

$$P\left(\theta \in \left[T_1\left(X\right), T_2\left(X\right)\right)\right] = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

- Статистики $T_1(X)$ и $T_2(X)$ именуются левой (нижней) и правой (верхней) границами доверительного интервала. Причем $P(T_2(X) \geq T_1(X)) = 1$.

Доверительные интервалы

Определение

- Рассмотрим выборку $X = (X_1, ..., X_n)$ из распределения с параметром θ.
- **Доверительным интервалом** с уровнем доверия $(1-\gamma)$, где $\gamma \in (0,1)$, называется интервал $[T_1(X), T_2(X)]$, такой, что при любом допустимом значении параметра θ:

$$P\left(\theta \in \left[T_1\left(X\right), T_2\left(X\right)\right)\right] = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

- Статистики $T_1(X)$ и $T_2(X)$ именуются левой (нижней) и правой (верхней) границами доверительного интервала. Причем $P(T_2(X) \geq T_1(X)) = 1$.
- Иногда говорят о $100(1-\gamma)$ процентном доверительном интервале, подразумевая уровень доверия $(1-\gamma)$.

- Рассмотрим выборку $X = (X_1, ..., X_n)$ из распределения с параметром θ.
- **Доверительным интервалом** с уровнем доверия $(1-\gamma)$, где $\gamma \in (0,1)$, называется интервал $[T_1(X), T_2(X)]$, такой, что при любом допустимом значении параметра θ:

$$P\left(\theta \in \left[T_1\left(X\right), T_2\left(X\right)\right)\right] = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

- Статистики $T_1(X)$ и $T_2(X)$ именуются левой (нижней) и правой (верхней) границами доверительного интервала. Причем $P(T_2(X) \geq T_1(X)) = 1$.
- Иногда говорят о $100(1-\gamma)$ процентном доверительном интервале, подразумевая уровень доверия $(1-\gamma)$.
  Пример: имеется выборка объема $n=1$ из равномерного распределения $X_1 \sim U(0,\alpha)$. Рассмотрим доверительный интервал с границами $T_1(X) = 2X_1$ и $T_2(X) = 5X_1$. Обратим внимание, что $X_1/\alpha \sim U(0,1)$ и рассчитаем уровень доверия:

- Рассмотрим выборку $X = (X_1, ..., X_n)$ из распределения с параметром θ.
- **Доверительным интервалом** с уровнем доверия $(1-\gamma)$, где $\gamma \in (0,1)$, называется интервал $[T_1(X), T_2(X)]$, такой, что при любом допустимом значении параметра θ:

$$P\left(\theta \in \left[T_1\left(X\right), T_2\left(X\right)\right)\right] = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

- Статистики $T_1(X)$ и $T_2(X)$ именуются левой (нижней) и правой (верхней) границами доверительного интервала. Причем $P(T_2(X) \geq T_1(X)) = 1$.
- Иногда говорят о $100(1-\gamma)$ процентном доверительном интервале, подразумевая уровень доверия $(1-\gamma)$.
  Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Рассмотрим доверительный интервал с границами $T_1(X) = 2X_1$ и $T_2(X) = 5X_1$. Обратим внимание, что $X_1/\alpha \sim U(0,1)$ и рассчитаем уровень доверия:

$$P\left(2X_1 \leq \alpha \leq 5X_1\right) = P\left(2 \leq \frac{\alpha}{X_1} \leq 5\right) = P\left(0.2 \leq \frac{X_1}{\alpha} \leq 0.5\right) = 0.5 - 0.2 = 0.3$$

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.
  Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Построим несколько доверительных интервалов с уровнем доверия 0.2.

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.
  Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Построим несколько доверительных интервалов с уровнем доверия 0.2.
  Симметричный: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = 2.5X_1$, поскольку:

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.
  Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Построим несколько доверительных интервалов с уровнем доверия 0.2.
  Симметричный: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = 2.5X_1$, поскольку:

$$P\left(5X_1/3 \leq \alpha \leq 2.5X_1\right) = P\left(0.4 \leq \frac{X_1}{\alpha} \leq 0.6\right) = 0.2$$

- Доверительный интервал является **симметричным**, если:
$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$
- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.
Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Построим несколько доверительных интервалов с уровнем доверия 0.2.
Симметричный: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = 2.5X_1$, поскольку:
$$P\left(5X_1/3 \leq \alpha \leq 2.5X_1\right) = P\left(0.4 \leq \frac{X_1}{\alpha} \leq 0.6\right) = 0.2$$

$$P\left(5X_1/3 > \alpha\right) = P\left(\frac{X_1}{\alpha} > 0.6\right) = 0.4 = P\left(\frac{X_1}{\alpha} < 0.4\right) = P\left(2.5X_1 < \alpha\right)$$

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.

  Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Построим несколько доверительных интервалов с уровнем доверия 0.2.

  Симметричный: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = 2.5X_1$, поскольку:

$$P\left(5X_1/3 \leq \alpha \leq 2.5X_1\right) = P\left(0.4 \leq \frac{X_1}{\alpha} \leq 0.6\right) = 0.2$$

  Правосторонний: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = \infty$, поскольку:

$$P\left(5X_1/3 > \alpha\right) = P\left(\frac{X_1}{\alpha} > 0.6\right) = 0.4 = P\left(\frac{X_1}{\alpha} < 0.4\right) = P\left(2.5X_1 < \alpha\right)$$

  Левосторонний: достаточно положить $T_1(X) = -\infty$ и $T_2(X) = 1.25$, поскольку:

- Доверительный интервал является **симметричным**, если:

$$P\left(T_2 < \theta\right) = P\left(T_1 > \theta\right)$$

- Доверительный интервал является **левосторонним** (верхним) если $T_1(X) = -\infty$.
- Доверительный интервал является **правосторонним** (нижним) если $T_2(X) = \infty$.
Пример: имеется выборка объема $n = 1$ из равномерного распределения $X_1 \sim U(0, \alpha)$. Построим несколько доверительных интервалов с уровнем доверия 0.2.
Симметричный: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = 2.5X_1$, поскольку:

$$P\left(5X_1/3 \leq \alpha \leq 2.5X_1\right) = P\left(0.4 \leq \frac{X_1}{\alpha} \leq 0.6\right) = 0.2$$

Правосторонний: достаточно положить $T_1(X) = 5X_1/3$ и $T_2(X) = \infty$, поскольку:

$$P\left(5X_1/3 > \alpha\right) = P\left(\frac{X_1}{\alpha} > 0.6\right) = 0.4 = P\left(\frac{X_1}{\alpha} < 0.4\right) = P\left(2.5X_1 < \alpha\right)$$

Левосторонний: достаточно положить $T_1(X) = -\infty$ и $T_2(X) = 1.25$, поскольку:

$$P\left(\alpha \leq 1.25X_1\right) = P\left(0.8 \leq \frac{X_1}{\alpha}\right) = 0.2$$

- Статистика $G(X, \theta)$ именуется **центральной**, если:

- Статистика $G(X,\theta)$ именуется **центральной**, если:
  - Распределение $G(X,\theta)$ остается неизменным при любом допустимом θ.

Алгоритм построения доверительного интервала с помощью центральной статистики

- Статистика $G(X, \theta)$ именуется **центральной**, если:
  - Распределение $G(X, \theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x, \theta)$ непрерывна и строго монотонна по θ.

- Статистика $G(X, \theta)$ именуется **центральной**, если:
  - Распределение $G(X, \theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x, \theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X, \theta)$ и для простоты предположим, что она имеет непрерывное распределение.

Алгоритм построения доверительного интервала с помощью центральной статистики

- Статистика $G(X, \theta)$ именуется **центральной**, если:
  - Распределение $G(X, \theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x, \theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X, \theta)$ и для простоты предположим, что она имеет непрерывное распределение.
- Для построения симметричного доверительного интервала уровня $(1 - \gamma)$ достаточно найти любую центральную статистику $G(X, \theta)$ и ее распределение, исходя из которого следует вычислить квантили $(\gamma/2)_G$ и $(1 - \gamma/2)_G$.

- Статистика $G(X, \theta)$ именуется **центральной**, если:
  - Распределение $G(X, \theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x, \theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X, \theta)$ и для простоты предположим, что она имеет непрерывное распределение.
- Для построения симметричного доверительного интервала уровня $(1-\gamma)$ достаточно найти любую центральную статистику $G(X, \theta)$ и ее распределение, исходя из которого следует вычислить квантили $(\gamma/2)_G$ и $(1-\gamma/2)_G$. Поскольку распределение $G(X, \theta)$ не зависит от θ, то от параметра также не будут зависеть и квантили. Поэтому, пользуясь непрерывностью и строгой монотонностью $G(X, \theta)$ можно найти такие $T_1(X)$ и $T_2(X)$, что:

Алгоритм построения доверительного интервала с помощью центральной статистики

- Статистика $G(X,\theta)$ именуется **центральной**, если:
  - Распределение $G(X,\theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x,\theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X,\theta)$ и для простоты предположим, что она имеет непрерывное распределение.
- Для построения симметричного доверительного интервала уровня $(1-\gamma)$ достаточно найти любую центральную статистику $G(X,\theta)$ и ее распределение, исходя из которого следует вычислить квантили $(\gamma/2)_G$ и $(1-\gamma/2)_G$. Поскольку распределение $G(X,\theta)$ не зависит от θ, то от параметра также не будут зависеть и квантили. Поэтому, пользуясь непрерывностью и строгой монотонностью $G(X,\theta)$ можно найти такие $T_1(X)$ и $T_2(X)$, что:

$$P\left(\frac{\gamma}{2}_G \leq G(X,\theta) \leq \left(1-\frac{\gamma}{2}\right)_G\right) = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

- Статистика $G(X,\theta)$ именуется **центральной**, если:
  - Распределение $G(X,\theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x,\theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X,\theta)$ и для простоты предположим, что она имеет непрерывное распределение.
- Для построения симметричного доверительного интервала уровня $(1-\gamma)$ достаточно найти любую центральную статистику $G(X,\theta)$ и ее распределение, исходя из которого следует вычислить квантили $(\gamma/2)_G$ и $(1-\gamma/2)_G$. Поскольку распределение $G(X,\theta)$ не зависит от θ, то от параметра также не будут зависеть и квантили. Поэтому, пользуясь непрерывностью и строгой монотонностью $G(X,\theta)$ можно найти такие $T_1(X)$ и $T_2(X)$, что:

$$P\left({\frac{\gamma}{2}}_G \leq G(X,\theta) \leq \left(1-\frac{\gamma}{2}\right)_G\right) = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

Пример: имеется выборка из равномерного распределения $U(0,\alpha)$. Статистика $G(X,\alpha)=\left(\frac{X_{(n)}}{\alpha}\right)^n \sim U(0,1)$, где $X_{(n)}=\max\left(X_1,...,X_n\right)$, является центральной, поскольку ее распределение не зависит от α (всегда стандартное равномерное) и функция $\left(\frac{\max(x_1,\ldots,x_n)}{\alpha}\right)^n$ непрерывна и строго монотонна по α.

- Статистика $G(X, \theta)$ именуется **центральной**, если:
  - Распределение $G(X, \theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x, \theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X, \theta)$ и для простоты предположим, что она имеет непрерывное распределение.
- Для построения симметричного доверительного интервала уровня $(1-\gamma)$ достаточно найти любую центральную статистику $G(X, \theta)$ и ее распределение, исходя из которого следует вычислить квантили $(\gamma/2)_G$ и $(1-\gamma/2)_G$. Поскольку распределение $G(X, \theta)$ не зависит от θ, то от параметра также не будут зависеть и квантили. Поэтому, пользуясь непрерывностью и строгой монотонностью $G(X, \theta)$ можно найти такие $T_1(X)$ и $T_2(X)$, что:

$$P\left(\frac{\gamma}{2}_G \leq G(X, \theta) \leq \left(1-\frac{\gamma}{2}\right)_G\right) = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

Пример: имеется выборка из равномерного распределения $U(0, \alpha)$. Статистика $G(X, \alpha) = \left(\frac{X_{(n)}}{\alpha}\right)^n \sim U(0, 1)$, где $X_{(n)} = \max(X_1, ..., X_n)$, является центральной, поскольку ее распределение не зависит от α (всегда стандартное равномерное) и функция $\left(\frac{\max(x_1, \ldots, x_n)}{\alpha}\right)^n$ непрерывна и строго монотонна по α. Для построения симметричного доверительного интервала уровня 0.2 рассмотрим квантили $0.1_G = 0.1$ и $0.9_G = 0.9$:

- Статистика $G(X,\theta)$ именуется **центральной**, если:
  - Распределение $G(X,\theta)$ остается неизменным при любом допустимом θ.
  - При любой реализации x функция $G(x,\theta)$ непрерывна и строго монотонна по θ.
- Через q_G обозначим квантиль уровня q центральной статистики $G(X,\theta)$ и для простоты предположим, что она имеет непрерывное распределение.
- Для построения симметричного доверительного интервала уровня $(1-\gamma)$ достаточно найти любую центральную статистику $G(X,\theta)$ и ее распределение, исходя из которого следует вычислить квантили $(\gamma/2)_G$ и $(1-\gamma/2)_G$. Поскольку распределение $G(X,\theta)$ не зависит от θ, то от параметра также не будут зависеть и квантили. Поэтому, пользуясь непрерывностью и строгой монотонностью $G(X,\theta)$ можно найти такие $T_1(X)$ и $T_2(X)$, что:

$$P\left(\frac{\gamma}{2}_G \leq G(X,\theta) \leq \left(1-\frac{\gamma}{2}\right)_G\right) = P\left(T_1(X) \leq \theta \leq T_2(X)\right) = 1-\gamma$$

Пример: имеется выборка из равномерного распределения $U(0,\alpha)$. Статистика $G(X,\alpha) = \left(\frac{X_{(n)}}{\alpha}\right)^n \sim U(0,1)$, где $X_{(n)} = \max\left(X_1,...,X_n\right)$, является центральной, поскольку ее распределение не зависит от α (всегда стандартное равномерное) и функция $\left(\frac{\max(x_1,\ldots,x_n)}{\alpha}\right)^n$ непрерывна и строго монотонна по α. Для построения симметричного доверительного интервала уровня 0.2 рассмотрим квантили $0.1_G = 0.1$ и $0.9_G = 0.9$:

$$P\left(0.1 \leq \left(\frac{X_{(n)}}{\alpha}\right)^n \leq 0.9\right) = P\left(\frac{X_{(n)}}{\sqrt[n]{0.9}} \leq \alpha \leq \frac{X_{(n)}}{\sqrt[n]{0.1}}\right) = 0.2$$

Доверительные интервалы
Проблема интерпретации

- Через $T_1(x)$ и $T_2(x)$ обозначим реализации границ доверительного интервала с уровнем доверия $(1-\gamma)$. Через $[T_1(x), T_2(x)]$ обозначим реализацию доверительного интервала.

Проблема интерпретации

- Через $T_1(x)$ и $T_2(x)$ обозначим реализации границ доверительного интервала с уровнем доверия $(1-\gamma)$. Через $[T_1(x), T_2(x)]$ обозначим реализацию доверительного интервала.
- Часто, получив реализации границ доверительного интервала для параметра θ исследователи утверждают, что параметр θ находится между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$.

- Через $T_1(x)$ и $T_2(x)$ обозначим реализации границ доверительного интервала с уровнем доверия $(1-\gamma)$. Через $[T_1(x), T_2(x)]$ обозначим реализацию доверительного интервала.
- Часто, получив реализации границ доверительного интервала для параметра θ исследователи утверждают, что параметр θ находится между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$. Однако, такой подход к интерпретации является ошибочным, поскольку $T_1(x)$ и $T_2(x)$ не являются случайными величинами и нет никакой гарантии, что при повторении эксперимента параметр θ будет оказываться между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$.

- Через $T_1(x)$ и $T_2(x)$ обозначим реализации границ доверительного интервала с уровнем доверия $(1-\gamma)$. Через $[T_1(x), T_2(x)]$ обозначим реализацию доверительного интервала.
- Часто, получив реализации границ доверительного интервала для параметра θ исследователи утверждают, что параметр θ находится между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$. Однако, такой подход к интерпретации является ошибочным, поскольку $T_1(x)$ и $T_2(x)$ не являются случайными величинами и нет никакой гарантии, что при повторении эксперимента параметр θ будет оказываться между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$.
- Фактически, $T_1(x)$ и $T_2(x)$ это реализации случайных величин, между которыми параметр θ мог оказаться с вероятностью $1-\gamma$.

- Через $T_1(x)$ и $T_2(x)$ обозначим реализации границ доверительного интервала с уровнем доверия $(1-\gamma)$. Через $[T_1(x), T_2(x)]$ обозначим реализацию доверительного интервала.
- Часто, получив реализации границ доверительного интервала для параметра θ исследователи утверждают, что параметр θ находится между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$. Однако, такой подход к интерпретации является ошибочным, поскольку $T_1(x)$ и $T_2(x)$ не являются случайными величинами и нет никакой гарантии, что при повторении эксперимента параметр θ будет оказываться между $T_1(x)$ и $T_2(x)$ с вероятностью $1-\gamma$.
- Фактически, $T_1(x)$ и $T_2(x)$ это реализации случайных величин, между которыми параметр θ мог оказаться с вероятностью $1-\gamma$.
- С практической точки зрения мы можем принять $[T_1(x), T_2(x)]$ как интервальную оценку (аппроксимацию) интервала, в котором истинное значение параметра лежит с вероятностью $1-\gamma$.

Математическое ожидание при известной дисперсии

- По выборке из нормального распределения $\mathcal{N}(\mu,\sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.

Доверительные интервалы для параметров нормального распределения

Математическое ожидание при известной дисперсии

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

- Обозначим через z_γ квантиль стандартного нормального распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $z_{1-\gamma/2} = -z_{\gamma/2}$:

$$P\left(z_{\gamma/2} \leq \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{1-\gamma/2}\right) = P\left(\overline{X}_n - z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X}_n + z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}}\right) = (1-\gamma)$$

Доверительные интервалы для параметров нормального распределения

Математическое ожидание при известной дисперсии

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

- Обозначим через z_γ квантиль стандартного нормального распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $z_{1-\gamma/2} = -z_{\gamma/2}$:

$$P\left(z_{\gamma/2} \leq \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{1-\gamma/2}\right) = P\left(\overline{X}_n - z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X}_n + z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}(\mu, 25)$ и с реализацией $x = (0, 2, 0, 6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9.

Доверительные интервалы для параметров нормального распределения

Математическое ожидание при известной дисперсии

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

- Обозначим через z_γ квантиль стандартного нормального распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $z_{1-\gamma/2} = -z_{\gamma/2}$:

$$P\left(z_{\gamma/2} \leq \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{1-\gamma/2}\right) = P\left(\overline{X}_n - z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X}_n + z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}(\mu, 25)$ и с реализацией $x = (0, 2, 0, 6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9. Поскольку $(1-\gamma) = 0.9$, то $\gamma = 0.1$, откуда по таблице стандартного нормального распределения находим $z_{1-0.1/2} = z_{0.95} \approx 1.645$. Кроме того $\overline{x}_4 = (0+2+0+6)/4 = 2$.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

- Обозначим через z_γ квантиль стандартного нормального распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $z_{1-\gamma/2} = -z_{\gamma/2}$:

$$P\left(z_{\gamma/2} \leq \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{1-\gamma/2}\right) = P\left(\overline{X}_n - z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \overline{X}_n + z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}(\mu, 25)$ и с реализацией $x = (0, 2, 0, 6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9. Поскольку $(1-\gamma) = 0.9$, то $\gamma = 0.1$, откуда по таблице стандартного нормального распределения находим $z_{1-0.1/2} = z_{0.95} \approx 1.645$. Кроме того $\overline{x}_4 = (0+2+0+6)/4 = 2$.

Реализация границ: $T_1(x) = 2 - 1.645\sqrt{\frac{25}{4}} \approx -2.1 \qquad T_2(x) = 2 + 1.645\sqrt{\frac{25}{4}} \approx 6.1$

- По выборке из нормального распределения $\mathcal{N}(\mu,\sigma^2)$ найдем симметричный доверительный интервал для μ предполагая, что параметр σ^2 известен.
- Центральная статистика будет иметь вид:

$$G(X,\mu)=\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\sim\mathcal{N}\left(0,1\right)$$

- Обозначим через z_γ квантиль стандартного нормального распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $z_{1-\gamma/2}=-z_{\gamma/2}$:

$$P\left(z_{\gamma/2}\leq\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\leq z_{1-\gamma/2}\right)=P\left(\overline{X}_n-z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}}\leq\mu\leq\overline{X}_n+z_{1-\gamma/2}\sqrt{\frac{\sigma^2}{n}}\right)=(1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}\left(\mu,25\right)$ и с реализацией $x=(0,2,0,6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9. Поскольку $(1-\gamma)=0.9$, то $\gamma=0.1$, откуда по таблице стандартного нормального распределения находим $z_{1-0.1/2}=z_{0.95}\approx1.645$. Кроме того $\overline{x}_4=(0+2+0+6)/4=2$.

Реализация границ: $T_1(x)=2-1.645\sqrt{\frac{25}{4}}\approx-2.1\qquad T_2(x)=2+1.645\sqrt{\frac{25}{4}}\approx6.1$

Реализация доверительного интервала: $[-2.1,6.1]$

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.

Дисперсия при известном математическом ожидании

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$P\left(\chi^2_{n,\gamma/2} \leq \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \leq \chi^2_{n,1-\gamma/2}\right) = P\left(\frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{n,1-\gamma/2}} \leq \sigma^2 \leq \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{n,\gamma/2}}\right) = (1-\gamma)$$

Доверительные интервалы для параметров нормального распределения

Дисперсия при известном математическом ожидании

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$P\left(\chi^2_{n,\gamma/2} \leq \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \leq \chi^2_{n,1-\gamma/2}\right) = P\left(\frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{n,1-\gamma/2}} \leq \sigma^2 \leq \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{n,\gamma/2}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}\left(1, \sigma^2\right)$ и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$P\left(\chi^2_{n,\gamma/2} \leq \frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \leq \chi^2_{n,1-\gamma/2}\right) = P\left(\frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{n,1-\gamma/2}} \leq \sigma^2 \leq \frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{n,\gamma/2}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}\left(1, \sigma^2\right)$ и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95. Поскольку $(1-\gamma) = 0.95$, то $\gamma = 0.05$, откуда по таблице хи-квадрат распределения с $n = 3$ степенями свободы находим $\chi^2_{3,0.05/2} \approx 0.216$ и $\chi^2_{3,1-0.05/2} \approx 9.348$. Отсюда получаем:

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$P\left(\chi^2_{n,\gamma/2} \leq \frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \leq \chi^2_{n,1-\gamma/2}\right) = P\left(\frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{n,1-\gamma/2}} \leq \sigma^2 \leq \frac{\sum\limits_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{n,\gamma/2}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}\left(1, \sigma^2\right)$ и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95. Поскольку $(1-\gamma) = 0.95$, то $\gamma = 0.05$, откуда по таблице хи-квадрат распределения с $n = 3$ степенями свободы находим $\chi^2_{3,0.05/2} \approx 0.216$ и $\chi^2_{3,1-0.05/2} \approx 9.348$. Отсюда получаем:

Реализация границ: $T_1(x) = \frac{(0-1)^2 + (1-1)^2 + (5-1)^2}{9.348} \approx 1.82 \quad T_2(x) = \frac{(0-1)^2 + (1-1)^2 + (5-1)^2}{0.216} \approx 78.7$

Дисперсия при известном математическом ожидании

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2 предполагая, что параметр μ известен.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$P\left(\chi^2_{n,\gamma/2} \leq \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \leq \chi^2_{n,1-\gamma/2}\right) = P\left(\frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{n,1-\gamma/2}} \leq \sigma^2 \leq \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{n,\gamma/2}}\right) = (1-\gamma)$$

Пример: по выборке из распределения $\mathcal{N}\left(1, \sigma^2\right)$ и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95. Поскольку $(1-\gamma) = 0.95$, то $\gamma = 0.05$, откуда по таблице хи-квадрат распределения с $n = 3$ степенями свободы находим $\chi^2_{3,0.05/2} \approx 0.216$ и $\chi^2_{3,1-0.05/2} \approx 9.348$. Отсюда получаем:

Реализация границ: $T_1(x) = \frac{(0-1)^2 + (1-1)^2 + (5-1)^2}{9.348} \approx 1.82 \quad T_2(x) = \frac{(0-1)^2 + (1-1)^2 + (5-1)^2}{0.216} \approx 78.7$

Реализация доверительного интервала: $[1.82, 78.7]$

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где σ_X^2 и σ_Y^2 известны. Построим доверительный интервал для разницы математических ожиданий $\mu_X - \mu_Y$.

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где σ_X^2 и σ_Y^2 известны. Построим доверительный интервал для разницы математических ожиданий $\mu_X - \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim \mathcal{N}(0, 1)$$

- Имеются две независимые выборки $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где σ_X^2 и σ_Y^2 известны. Построим доверительный интервал для разницы математических ожиданий $\mu_X - \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim \mathcal{N}(0,1)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_m) - z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}, (\overline{X}_n - \overline{Y}_m) + z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где σ_X^2 и σ_Y^2 известны. Построим доверительный интервал для разницы математических ожиданий $\mu_X - \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim \mathcal{N}\left(0, 1\right)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_m) - z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}, (\overline{X}_n - \overline{Y}_m) + z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(\mu_X, 1\right)$ и $\mathcal{N}\left(\mu_Y, 9\right)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 99%-го доверительного интервала.

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где σ_X^2 и σ_Y^2 известны. Построим доверительный интервал для разницы математических ожиданий $\mu_X - \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim \mathcal{N}(0, 1)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_m) - z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}, (\overline{X}_n - \overline{Y}_m) + z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}(\mu_X, 1)$ и $\mathcal{N}(\mu_Y, 9)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 99%-го доверительного интервала. Поскольку $n = 2$, $m = 3$, $\overline{x}_2 = 2$, $\overline{y}_3 = 1$, $\sigma_X^2 = 1$, $\sigma_Y^2 = 3$ и $z_{0.995} \approx 2.58$, то искомая реализация равняется:

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где σ_X^2 и σ_Y^2 известны. Построим доверительный интервал для разницы математических ожиданий $\mu_X - \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim \mathcal{N}\left(0, 1\right)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_m) - z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}, (\overline{X}_n - \overline{Y}_m) + z_{1-\frac{\gamma}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(\mu_X, 1\right)$ и $\mathcal{N}\left(\mu_Y, 9\right)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 99%-го доверительного интервала. Поскольку $n = 2$, $m = 3$, $\overline{x}_2 = 2$, $\overline{y}_3 = 1$, $\sigma_X^2 = 1$, $\sigma_Y^2 = 3$ и $z_{0.995} \approx 2.58$, то искомая реализация равняется:

$$\left[(2-1) - 2.58\sqrt{\frac{1}{2} + \frac{9}{3}}, (2-1) + 2.58\sqrt{\frac{1}{2} + \frac{9}{3}}\right] \approx [-3.83, 5.83]$$

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где μ_X и μ_Y известны. Построим доверительный интервал для отношения дисперсий σ_Y^2/σ_X^2.

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где μ_X и μ_Y известны. Построим доверительный интервал для отношения дисперсий σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i-\mu_X\right)^2}{\sum_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n,m)$$

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где μ_X и μ_Y известны. Построим доверительный интервал для отношения дисперсий σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i-\mu_X\right)^2}{\sum_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n,m)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\mu_X\right)^2}\frac{n}{m}F_{n,m}^{(\gamma/2)}, \frac{\sum\limits_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\mu_X\right)^2}\frac{n}{m}F_{n,m}^{(1-\gamma/2)}\right]$$

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где μ_X и μ_Y известны. Построим доверительный интервал для отношения дисперсий σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i - \mu_X\right)^2}{\sum_{i=1}^{m}\left(Y_i - \mu_Y\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n, m)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i - \mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \mu_X\right)^2}\frac{n}{m}F_{n,m}^{(\gamma/2)}, \frac{\sum\limits_{i=1}^{m}\left(Y_i - \mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \mu_X\right)^2}\frac{n}{m}F_{n,m}^{(1-\gamma/2)}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(0, \sigma_X^2\right)$ и $\mathcal{N}\left(1, \sigma_Y^2\right)$.

Доверительные интервалы для параметров нормального распределения

Отношение дисперсий при известных математических ожиданиях

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно, где μ_X и μ_Y известны. Построим доверительный интервал для отношения дисперсий σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i-\mu_X\right)^2}{\sum_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n,m)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\mu_X\right)^2}\frac{n}{m}F_{n,m}^{(\gamma/2)}, \frac{\sum\limits_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\mu_X\right)^2}\frac{n}{m}F_{n,m}^{(1-\gamma/2)}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(0, \sigma_X^2\right)$ и $\mathcal{N}\left(1, \sigma_Y^2\right)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 90%-го доверительного интервала для σ_X^2/σ_Y^2:

- Имеются две независимые выборки $X=(X_1,...,X_n)$ и $Y=(Y_1,...,Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X,\sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y,\sigma_Y^2\right)$ соответственно, где μ_X и μ_Y известны. Построим доверительный интервал для отношения дисперсий σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i-\mu_X\right)^2}{\sum_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n}\sim F(n,m)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\mu_X\right)^2}\frac{n}{m}F_{n,m}^{(\gamma/2)},\ \frac{\sum\limits_{i=1}^{m}\left(Y_i-\mu_Y\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\mu_X\right)^2}\frac{n}{m}F_{n,m}^{(1-\gamma/2)}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(0,\sigma_X^2\right)$ и $\mathcal{N}\left(1,\sigma_Y^2\right)$. Реализации выборок имеют вид $x=(1,3)$ и $y=(1,0,2)$. Найдем реализацию 90%-го доверительного интервала для σ_X^2/σ_Y^2:

$$\left[\frac{(1-1)^2+(0-1)^2+(2-1)^2}{(1-0)^2+(3-0)^2}\frac{2}{3}0.052,\ \frac{(1-1)^2+(0-1)^2+(2-1)^2}{(1-0)^2+(3-0)^2}\frac{2}{3}9.55\right]\approx[0.007,1.273]$$

Доверительные интервалы для параметров нормального распределения

Дополнительный пример

Доходы программистов и аналитиков (в тысячах рублей) хорошо описываются нормальным распределением. Вы случайным образом опросили 3 программиста и 2 аналитика. Доходы программистов оказались равны 100, 80 и 120, а аналитиков – 90 и 130. Считая доходы опрошенных индивидов независимыми постройте 90% доверительный интервал для:

- Ожидаемого дохода случайно взятого программиста, если дисперсия его дохода равняется 225.
- Дисперсии дохода случайно взятого программиста, если его ожидаемый доход равен 120.
- Ожидаемой разницы в доходах случайно взятых программиста и аналитика, если их дисперсии равны и составляют 225.
- Отношения дисперсий доходов случайно взятого программиста к дисперсии дохода случайного взятого аналитика, если их ожидаемые доходы равны и составляют 100.

Доверительные интервалы для параметров нормального распределения
Дополнительный пример

Доходы программистов и аналитиков (в тысячах рублей) хорошо описываются нормальным распределением. Вы случайным образом опросили 3 программиста и 2 аналитика. Доходы программистов оказались равны 100, 80 и 120, а аналитиков – 90 и 130. Считая доходы опрошенных индивидов независимыми постройте 90% доверительный интервал для:

- Ожидаемого дохода случайно взятого программиста, если дисперсия его дохода равняется 225.
- Дисперсии дохода случайно взятого программиста, если его ожидаемый доход равен 120.
- Ожидаемой разницы в доходах случайно взятых программиста и аналитика, если их дисперсии равны и составляют 225.
- Отношения дисперсий доходов случайно взятого программиста к дисперсии дохода случайного взятого аналитика, если их ожидаемые доходы равны и составляют 100.

Решение:

- $[(100+80+120)/3 - 1.645 \times \sqrt{225/3}, (100+80+120)/3 + 1.645 \times \sqrt{225/3}] \approx [85.75, 114.25]$

Доходы программистов и аналитиков (в тысячах рублей) хорошо описываются нормальным распределением. Вы случайным образом опросили 3 программиста и 2 аналитика. Доходы программистов оказались равны 100, 80 и 120, а аналитиков – 90 и 130. Считая доходы опрошенных индивидов независимыми постройте 90% доверительный интервал для:

- Ожидаемого дохода случайно взятого программиста, если дисперсия его дохода равняется 225.
- Дисперсии дохода случайно взятого программиста, если его ожидаемый доход равен 120.
- Ожидаемой разницы в доходах случайно взятых программиста и аналитика, если их дисперсии равны и составляют 225.
- Отношения дисперсий доходов случайно взятого программиста к дисперсии дохода случайного взятого аналитика, если их ожидаемые доходы равны и составляют 100.

Решение:

- $[(100+80+120)/3-1.645\times\sqrt{225/3},(100+80+120)/3+1.645\times\sqrt{225/3}]\approx[85.75,114.25]$
- $\left[\frac{(100-120)^2+(80-120)^2+(120-120)^2}{7.81},\frac{(100-120)^2+(80-120)^2+(120-120)^2}{0.35}\right]\approx[256.08,5714.29]$

Доходы программистов и аналитиков (в тысячах рублей) хорошо описываются нормальным распределением. Вы случайным образом опросили 3 программиста и 2 аналитика. Доходы программистов оказались равны 100, 80 и 120, а аналитиков – 90 и 130. Считая доходы опрошенных индивидов независимыми постройте 90% доверительный интервал для:

- Ожидаемого дохода случайно взятого программиста, если дисперсия его дохода равняется 225.
- Дисперсии дохода случайно взятого программиста, если его ожидаемый доход равен 120.
- Ожидаемой разницы в доходах случайно взятых программиста и аналитика, если их дисперсии равны и составляют 225.
- Отношения дисперсий доходов случайно взятого программиста к дисперсии дохода случайного взятого аналитика, если их ожидаемые доходы равны и составляют 100.

Решение:

- $[(100+80+120)/3-1.645\times\sqrt{225/3},(100+80+120)/3+1.645\times\sqrt{225/3}]\approx[85.75,114.25]$
- $\left[\frac{(100-120)^2+(80-120)^2+(120-120)^2}{7.81},\frac{(100-120)^2+(80-120)^2+(120-120)^2}{0.35}\right]\approx[256.08,5714.29]$
- $[((100+80+120)/3-(90+130)/2)-1.645\sqrt{\frac{225}{3}+\frac{225}{2}},...+...]\approx[-32.5,12.5]$

Доверительные интервалы для параметров нормального распределения

Дополнительный пример

Доходы программистов и аналитиков (в тысячах рублей) хорошо описываются нормальным распределением. Вы случайным образом опросили 3 программиста и 2 аналитика. Доходы программистов оказались равны 100, 80 и 120, а аналитиков – 90 и 130. Считая доходы опрошенных индивидов независимыми постройте 90% доверительный интервал для:

- Ожидаемого дохода случайно взятого программиста, если дисперсия его дохода равняется 225.
- Дисперсии дохода случайно взятого программиста, если его ожидаемый доход равен 120.
- Ожидаемой разницы в доходах случайно взятых программиста и аналитика, если их дисперсии равны и составляют 225.
- Отношения дисперсий доходов случайно взятого программиста к дисперсии дохода случайного взятого аналитика, если их ожидаемые доходы равны и составляют 100.

Решение:

- $[(100+80+120)/3 - 1.645 \times \sqrt{225/3}, (100+80+120)/3 + 1.645 \times \sqrt{225/3}] \approx [85.75, 114.25]$
- $\left[\frac{(100-120)^2+(80-120)^2+(120-120)^2}{7.81}, \frac{(100-120)^2+(80-120)^2+(120-120)^2}{0.35}\right] \approx [256.08, 5714.29]$
- $[((100+80+120)/3 - (90+130)/2) - 1.645\sqrt{\frac{225}{3}+\frac{225}{2}}, ... + ...] \approx [-32.5, 12.5]$
- $\left[\frac{(100-100)^2+(80-100)^2+(120-100)^2}{(90-100)^2+(130-100)^2}\frac{2}{3}0.052, \frac{(100-100)^2+(80-100)^2+(120-100)^2}{(90-100)^2+(130-100)^2}\frac{2}{3}9.55\right] \approx [0.028, 5.093]$

- Вспомним формулу исправленной выборочной дисперсии:

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

Теорема Фишера

Формулировка и применение

- Вспомним формулу исправленной выборочной дисперсии:

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

- Согласно теореме Фишера:

$$\frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2(n-1)$$

Кроме того, $\hat{\sigma}_n^2$ и $\overline{X}_n$ независимы.

- Вспомним формулу исправленной выборочной дисперсии:

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

- Согласно теореме Фишера:

$$\frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2(n-1)$$

Кроме того, $\hat{\sigma}_n^2$ и $\overline{X}_n$ независимы.

- При помощи теоремы Фишера покажем следующий результат:

$$\frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} = \frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} * \frac{\left(\frac{\sqrt{n-1}}{\sigma}\right)}{\left(\frac{\sqrt{n-1}}{\sigma}\right)} = \frac{\frac{\overline{X}_n-\mu}{\sigma}\sqrt{n-1}}{\frac{\hat{\sigma}_n\sqrt{n-1}}{\sigma\sqrt{n}}} = \frac{\frac{\overline{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}}}{\frac{1}{\sqrt{n-1}}\frac{\hat{\sigma}_n\sqrt{n-1}}{\sigma}} = \frac{\frac{\overline{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{1}{n-1}\frac{\hat{\sigma}_n^2(n-1)}{\sigma^2}}} \sim t(n-1)$$

Мы получили распределение Стьюдента, поскольку, в силу теоремы Фишера, в последнем выражении числитель и знаменатель независимы, причем числитель имеет стандартное нормальное распределение, а выражение под корнем в знаменателе, делимое на $n-1$, – Хи-квадрат распределение с $(n-1)$-й степенью свободы.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} \sim t(n-1)$$

Доверительные интервалы для параметров нормального распределения
Математическое ожидание при неизвестной дисперсии

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} \sim t(n-1)$$

- Обозначим через $t_{n,\gamma}$ квантиль уровня γ распределения Стьюдента с n степенями свободы и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $t_{n,1-\gamma/2} = -t_{n,\gamma}$:

$$\left[\overline{X}_n - t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}, \overline{X}_n + t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}\right]$$

Доверительные интервалы для параметров нормального распределения

Математическое ожидание при неизвестной дисперсии

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} \sim t(n-1)$$

- Обозначим через $t_{n,\gamma}$ квантиль уровня γ распределения Стьюдента с n степенями свободы и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $t_{n,1-\gamma/2} = -t_{n,\gamma}$:

$$\left[\overline{X}_n - t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}, \overline{X}_n + t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}\right]$$

Пример: по выборке из нормального распределения и с реализацией $x = (0, 2, 0, 6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} \sim t(n-1)$$

- Обозначим через $t_{n,\gamma}$ квантиль уровня γ распределения Стьюдента с n степенями свободы и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $t_{n,1-\gamma/2} = -t_{n,\gamma}$:

$$\left[\overline{X}_n - t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}, \overline{X}_n + t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}\right]$$

Пример: по выборке из нормального распределения и с реализацией $x = (0, 2, 0, 6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9. По таблице распределения Стьюдента находим $t_{4-1,0.95} \approx 2.35$. Кроме того $\overline{x}_4 = (0+2+0+6)/4 = 2$ и $\hat{\sigma}_4^2(x) = \left((0-2)^2 + (2-2)^2 + (0-2)^2 + (6-2)^2\right)/(4-1) = 8$.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для μ.
- Центральная статистика будет иметь вид:

$$G(X, \mu) = \frac{\overline{X}_n - \mu}{\frac{\hat{\sigma}_n}{\sqrt{n}}} \sim t(n-1)$$

- Обозначим через $t_{n,\gamma}$ квантиль уровня γ распределения Стьюдента с n степенями свободы и запишем выражение для доверительного интервала уровня $(1-\gamma)$. При этом воспользуемся тем, что $t_{n,1-\gamma/2} = -t_{n,\gamma}$:

$$\left[\overline{X}_n - t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}, \overline{X}_n + t_{n-1,1-\gamma/2}\sqrt{\frac{\hat{\sigma}_n^2}{n}}\right]$$

Пример: по выборке из нормального распределения и с реализацией $x = (0, 2, 0, 6)$ найдем реализацию доверительного интервала с уровнем доверия 0.9. По таблице распределения Стьюдента находим $t_{4-1,0.95} \approx 2.35$. Кроме того $\overline{x}_4 = (0+2+0+6)/4 = 2$ и $\hat{\sigma}_4^2(x) = \left((0-2)^2 + (2-2)^2 + (0-2)^2 + (6-2)^2\right)/(4-1) = 8$.

$$\left[2 - 2.35\sqrt{\frac{8}{4}}, 2 + 2.35\sqrt{\frac{8}{4}}\right] \approx [-1.32, 5.32]$$

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2\,(n-1)$$

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2\,(n-1)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$\left[\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,\gamma/2}}\right] = \left[\frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,\gamma/2}}\right]$$

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2(n-1)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$\left[\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,\gamma/2}}\right] = \left[\frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,\gamma/2}}\right]$$

Пример: по выборке из нормального распределения и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95.

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2(n-1)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$\left[\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,\gamma/2}}\right] = \left[\frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,\gamma/2}}\right]$$

Пример: по выборке из нормального распределения и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95. Поскольку $(1-\gamma) = 0.95$, то $\gamma = 0.05$, откуда по таблице хи-квадрат распределения с $n = 3 - 1 = 2$ степенями свободы находим $\chi^2_{2,0.05/2} \approx 0.05$ и $\chi^2_{2,1-0.05/2} \approx 7.38$. Кроме того, $\overline{x}_3 = (0+1+5)/3 = 2$. Отсюда получаем искомую реализацию:

- По выборке из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ найдем симметричный доверительный интервал для σ^2.
- Центральная статистика будет иметь вид:

$$G(X, \sigma^2) = \frac{(n-1)\hat{\sigma}_n^2}{\sigma^2} \sim \chi^2(n-1)$$

- Обозначим через $\chi^2_{n,\gamma}$ квантиль Хи-квадрат распределения уровня γ и запишем выражение для доверительного интервала уровня $(1-\gamma)$:

$$\left[\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\chi^2_{n-1,\gamma/2}}\right] = \left[\frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{(n-1)\hat{\sigma}_n^2}{\chi^2_{n-1,\gamma/2}}\right]$$

Пример: по выборке из нормального распределения и с реализацией $x = (0, 1, 5)$ найдем реализацию доверительного интервала с уровнем доверия 0.95. Поскольку $(1-\gamma) = 0.95$, то $\gamma = 0.05$, откуда по таблице хи-квадрат распределения с $n = 3 - 1 = 2$ степенями свободы находим $\chi^2_{2,0.05/2} \approx 0.05$ и $\chi^2_{2,1-0.05/2} \approx 7.38$. Кроме того, $\overline{x}_3 = (0+1+5)/3 = 2$. Отсюда получаем искомую реализацию:

$$\left[\frac{(0-2)^2 + (1-2)^2 + (5-2)^2}{7.38}, \frac{(0-2)^2 + (1-2)^2 + (5-2)^2}{0.05}\right] \approx [1.9, 280]$$

Доверительные интервалы для параметров нормального распределения

Разница математических ожиданий при равных по объему выборках

- Имеются две равные по объему выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_n)$, такие, что $X - Y$ является выборкой из нормального распределения. Построим доверительный интервал для $\mu_X - \mu_Y$, где $E(X_1) = \mu_X$ и $E(Y_1) = \mu_Y$.

.

Доверительные интервалы для параметров нормального распределения

Разница математических ожиданий при равных по объему выборках

- Имеются две равные по объему выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_n)$, такие, что $X - Y$ является выборкой из нормального распределения. Построим доверительный интервал для $\mu_X - \mu_Y$, где $E(X_1) = \mu_X$ и $E(Y_1) = \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_n\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}} \sim t\left(n-1\right), \text{ где } \hat{\sigma}^2_{X-Y} \text{ считается по выборке } (X - Y)$$

.

- Имеются две равные по объему выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_n)$, такие, что $X - Y$ является выборкой из нормального распределения. Построим доверительный интервал для $\mu_X - \mu_Y$, где $E(X_1) = \mu_X$ и $E(Y_1) = \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_n\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}} \sim t\,(n-1)\text{, где } \hat{\sigma}^2_{X-Y} \text{ считается по выборке } (X - Y)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_n) - t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}, (\overline{X}_n - \overline{Y}_n) + t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}\right]$$

.

- Имеются две равные по объему выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_n)$, такие, что $X - Y$ является выборкой из нормального распределения. Построим доверительный интервал для $\mu_X - \mu_Y$, где $E(X_1) = \mu_X$ и $E(Y_1) = \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_n\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}} \sim t\,(n-1)\,,\text{ где } \hat{\sigma}^2_{X-Y} \text{ считается по выборке } (X - Y)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_n) - t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}, (\overline{X}_n - \overline{Y}_n) + t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}\right]$$

Пример: имеются выборки X и Y, такие, что $(X - Y)$ имеет нормальное распределение, причем $x = (0, 6)$ и $y = (0, 2)$. Найдем реализацию 99%-го доверительного интервала для разницы математических ожиданий.

- Имеются две равные по объему выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_n)$, такие, что $X - Y$ является выборкой из нормального распределения. Построим доверительный интервал для $\mu_X - \mu_Y$, где $E(X_1) = \mu_X$ и $E(Y_1) = \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_n\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}} \sim t\left(n-1\right), \text{ где } \hat{\sigma}^2_{X-Y} \text{ считается по выборке } (X-Y)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_n) - t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}, (\overline{X}_n - \overline{Y}_n) + t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}\right]$$

Пример: имеются выборки X и Y, такие, что $(X - Y)$ имеет нормальное распределение, причем $x = (0, 6)$ и $y = (0, 2)$. Найдем реализацию 99%-го доверительного интервала для разницы математических ожиданий. Поскольку $\overline{x}_2 = 3$, $\overline{y}_2 = 1$, $x - y = (0, 4)$, $\overline{(x-y)}_2 = 2$, $\hat{\sigma}^2_2 = \left((0-2)^2 + (4-2)^2\right)/(2-1) = 8$, $t_{5-1,0.995} \approx 4.6$, то:

- Имеются две равные по объему выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_n)$, такие, что $X - Y$ является выборкой из нормального распределения. Построим доверительный интервал для $\mu_X - \mu_Y$, где $E(X_1) = \mu_X$ и $E(Y_1) = \mu_Y$.
- Центральная статистика будет иметь вид:

$$\frac{\left(\overline{X}_n - \overline{Y}_n\right) - (\mu_X - \mu_Y)}{\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}} \sim t\,(n-1)\,,\text{ где } \hat{\sigma}^2_{X-Y} \text{ считается по выборке } (X - Y)$$

- Нетрудно показать, что $100(1-\gamma)$ процентный доверительный интервал для $\mu_X - \mu_Y$ будет иметь вид:

$$\left[(\overline{X}_n - \overline{Y}_n) - t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}, (\overline{X}_n - \overline{Y}_n) + t_{n-1,1-\frac{\gamma}{2}}\sqrt{\frac{\hat{\sigma}^2_{X-Y}}{n}}\right]$$

Пример: имеются выборки X и Y, такие, что $(X - Y)$ имеет нормальное распределение, причем $x = (0, 6)$ и $y = (0, 2)$. Найдем реализацию 99%-го доверительного интервала для разницы математических ожиданий. Поскольку $\overline{x}_2 = 3$, $\overline{y}_2 = 1$, $x - y = (0, 4)$, $\overline{(x-y)}_2 = 2$, $\hat{\sigma}^2_2 = \left((0-2)^2 + (4-2)^2\right)/(2-1) = 8$, $t_{5-1,0.995} \approx 4.6$, то:

$$\left[2 - 4.6\sqrt{\frac{8}{2}}, 2 + 4.6\sqrt{\frac{8}{2}}\right] \approx [-7.2, 11.2]$$

Доверительные интервалы для параметров нормального распределения

Отношение дисперсий при неизвестных математических ожиданиях

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно. Построим доверительный интервал для σ_Y^2/σ_X^2.

.

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно. Построим доверительный интервал для σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\sum_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n-1, m-1)$$

.

- Имеются две независимые выборки $X=(X_1,...,X_n)$ и $Y=(Y_1,...,Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X,\sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y,\sigma_Y^2\right)$ соответственно. Построим доверительный интервал для σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2}{\sum_{i=1}^{m}\left(Y_i-\overline{Y}_m\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n}\sim F(n-1,m-1)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i-\overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(\gamma/2)},\frac{\sum\limits_{i=1}^{m}\left(Y_i-\overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(1-\gamma/2)}\right]=\left[\frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(\gamma/2)},\frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(1-\gamma/2)}\right]$$

.

Доверительные интервалы для параметров нормального распределения

Отношение дисперсий при неизвестных математических ожиданиях

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно. Построим доверительный интервал для σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\sum_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n-1, m-1)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(\gamma/2)}, \frac{\sum\limits_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(1-\gamma/2)}\right] = \left[\frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(\gamma/2)}, \frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(1-\gamma/2)}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 90%-го доверительного интервала для σ_Y^2/σ_X^2.

Доверительные интервалы для параметров нормального распределения

Отношение дисперсий при неизвестных математических ожиданиях

- Имеются две независимые выборки $X = (X_1, \dots, X_n)$ и $Y = (Y_1, \dots, Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно. Построим доверительный интервал для σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\sum_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n-1, m-1)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(\gamma/2)}, \frac{\sum\limits_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(1-\gamma/2)}\right] = \left[\frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(\gamma/2)}, \frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(1-\gamma/2)}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 90%-го доверительного интервала для σ_Y^2/σ_X^2. Учитывая, что $\overline{x}_2 = (1+3)/2 = 2$ и $\overline{y}_3 = (1+0+2)/3 = 1$, получаем:

- Имеются две независимые выборки $X = (X_1, ..., X_n)$ и $Y = (Y_1, ..., Y_m)$ из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$ соответственно. Построим доверительный интервал для σ_Y^2/σ_X^2.
- Центральная статистика будет иметь вид:

$$\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\sum_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}\frac{\sigma_Y^2}{\sigma_X^2}\frac{m}{n} \sim F(n-1, m-1)$$

- Обозначим через $F_{n,m}^{(\gamma)}$ квантиль уровня γ распределения Фишера с n и m степенями свободы, откуда получаем $100(1-\gamma)$ процентный доверительный интервал для σ_Y^2/σ_X^2:

$$\left[\frac{\sum\limits_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(\gamma/2)}, \frac{\sum\limits_{i=1}^{m}\left(Y_i - \overline{Y}_m\right)^2}{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}\frac{n-1}{m-1}F_{n-1,m-1}^{(1-\gamma/2)}\right] = \left[\frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(\gamma/2)}, \frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2}F_{n-1,m-1}^{(1-\gamma/2)}\right]$$

Пример: имеются независимые выборки из нормальных распределений $\mathcal{N}\left(\mu_X, \sigma_X^2\right)$ и $\mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$. Реализации выборок имеют вид $x = (1, 3)$ и $y = (1, 0, 2)$. Найдем реализацию 90%-го доверительного интервала для σ_Y^2/σ_X^2. Учитывая, что $\overline{x}_2 = (1+3)/2 = 2$ и $\overline{y}_3 = (1+0+2)/3 = 1$, получаем:

$$\left[\frac{(1-1)^2+(0-1)^2+(2-1)^2}{(1-2)^2+(3-2)^2}\frac{2-1}{3-1}0.005, \frac{(1-1)^2+(0-1)^2+(2-1)^2}{(1-2)^2+(3-2)^2}\frac{2-1}{3-1}18.5\right] \approx [0.0025, 9.25]$$

- Если в задаче идет речь о построении доверительного интервала с использованием выборки из нормального распределения, то для выбора нужной формулы достаточно ответить на несколько простых вопросов.

- Если в задаче идет речь о построении доверительного интервала с использованием выборки из нормального распределения, то для выбора нужной формулы достаточно ответить на несколько простых вопросов.
- Известен ли один из двух параметров нормального распределения: μ или σ^2?

- Если в задаче идет речь о построении доверительного интервала с использованием выборки из нормального распределения, то для выбора нужной формулы достаточно ответить на несколько простых вопросов.
- Известен ли один из двух параметров нормального распределения: μ или σ^2?
- Доверительный интервал касается математического ожидания, дисперсии, разницы математических ожиданий или отношения дисперсий?