

ТЕОРИЯ ВЕРОЯТНОСТИ



Лекция 12

Точечные оценки

доцент:

Колосько Анатолий Григорьевич

(AGKOLOSKO@MAIL.RU)

Необходимость оценки распределения

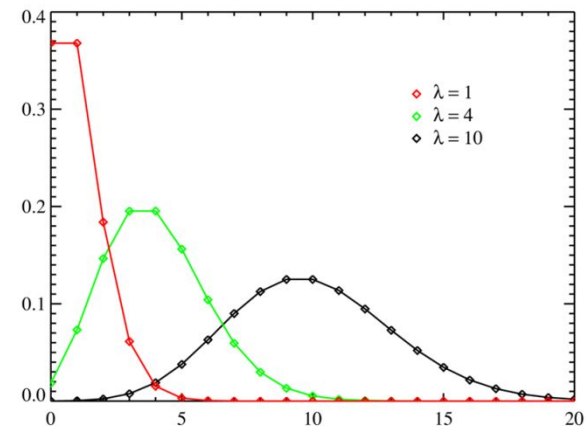
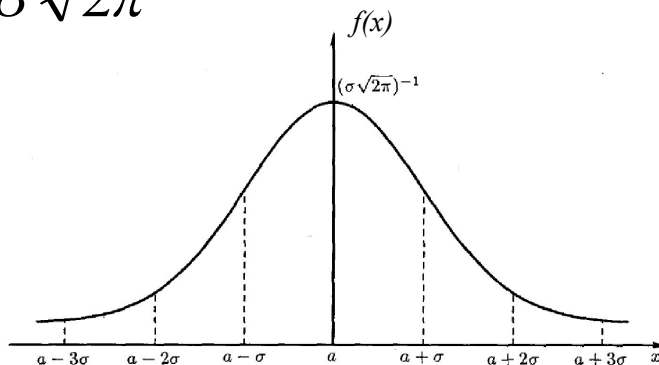
Пусть требуется изучить количественный признак генеральной совокупности. Допустим, что из теоретических соображений удалось установить, какое именно распределение имеет признак.

если наперед известно, что изучаемый признак распределен в генеральной совокупности нормально, то необходимо оценить (приблизительно найти) математическое ожидание и среднее квадратическое отклонение,

признак имеет, например, распределение Пуассона, то необходимо оценить параметр λ , которым это распределение определяется.

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

$$P_n(m) = \frac{\mu^m}{m!} e^{-\mu}$$



Статистическая оценка

Выборку объёма n : $V = \{x_1, x_2, \dots, x_n\}$ можно назвать n -мерной случайной величиной. Любая функция от этой величины называется **статистикой**.

Статистической оценкой неизвестного параметра теоретического распределения называют функцию от наблюдаемых случайных величин (т.е. статистику), которая его описывает.

Оценку, которая определяется одним числом, называют **точечной**.

Например: для оценки математического ожидания нормального распределения служит функция (среднее арифметическое наблюдаемых значений признака)

$$\bar{X} = (X_1 + X_2 + \dots + X_n)/n.$$

Чтобы статистические оценки давали хорошие приближения оцениваемых параметров, они должны удовлетворять определённым требованиям, которые мы рассмотрим далее.

Оценка как случайная величина

Обозначим Θ^* – статистическую оценку неизвестного параметра Θ_T теоретического распределения $f(x)$, описывающего случайную величину X .

Допустим, по выборке объёма n найдена оценка Θ_{1}^* .

Повторим опыт – из генеральной совокупности извлечём ещё одну выборку объёма n и получим оценку Θ_{2}^* .

Извлекая выборку многократно, получим набор РАЗЛИЧНЫХ статистических оценок $\Theta_{1}^*, \Theta_{2}^*, \dots, \Theta_{k}^*$.

Здесь мы **можем рассматривать Θ^* как случайную величину**, а множество $\{ \Theta_{1}^*, \dots, \Theta_{k}^* \}$ тогда будет набором её возможных значений.

Для этой случайной величины можно ввести те же самые вероятностные параметры, что и для обычной случайной величины X : среднее, дисперсию... при этом Θ сама, как параметр, может быть средней (например, для $Norm(a, \sigma)$).

Требования к оценкам статистических параметров

1. **Состоятельная статистическая оценка** – это оценка, которая при увеличении

объёма выборки $n \rightarrow \infty$ стремится по вероятности к оцениваемому параметру.

Требование состоятельности предъявляется при рассмотрении больших выборок. $P(|\Theta_n^* - \Theta_T| > \varepsilon) \rightarrow 0$ при $n \rightarrow \infty$ для $\forall \varepsilon > 0$

2. **Несмещённая статистическая оценка** – это оценка, у которой мат. ожидание $M[\Theta_n^*] = \Theta_T$ при $\forall n$ равно оцениваемому параметру при любом объёме выборки:

$$n \rightarrow \infty \quad D(\Theta_n^*) \rightarrow 0$$

Соблюдение этого требования гарантирует отсутствие систематических ошибок.

Если у несмещённой Θ при $n \rightarrow \infty$, то оценка ещё и состоятельная.

Смещённая статистическая оценка – это оценка, у которой наоборот :

$$M(\Theta^*) \neq \Theta_T \quad D[\Theta_n^*] = \min D(\text{различных } \Theta) \text{ при заданном } n$$

3. **Эффективная статистическая оценка** – это оценка, которая при

Генеральная средняя

Генеральной средней \bar{x}_r называют среднее арифметическое значений признака генеральной совокупности.

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты N_1, N_2, \dots, N_k , причем $N_1 + N_2 + \dots + N_k = N$, то

$$\bar{x}_r = (x_1 N_1 + x_2 N_2 + \dots + x_k N_k) / N,$$

Итак, если рассматривать обследуемый признак X генеральной совокупности как случайную величину, то математическое ожидание признака равно генеральной средней этого признака:

$$M(X) = \bar{x}_r.$$

Отклонением называют разность $x_i - \bar{x}$ между значением признака и общей средней.

Сумма этих отклонений должна быть равна 0, что следует из определения среднего:

$$\sum n_i (x_i - \bar{x}) = 0.$$

Генеральная средняя = мат. ожидание!

Математическое ожидание – это число: $M(X) = \sum_{i=1}^n (x_i \cdot p_i)$

Пусть в генеральной совокупности варианты-значения встречаются:

x_1 - n_1 раз, x_2 - n_2 раз ... x_m - n_m раз.

Среднее арифметическое всех результатов будет равно:

$$\langle X \rangle = (x_1 n_1 + x_2 n_2 + \dots + x_m n_m) \cdot \frac{1}{n} = x_1 \cdot \frac{n_1}{n} + x_2 \cdot \frac{n_2}{n} + \dots + x_m \cdot \frac{n_m}{n}$$

Так как при выборке число X мы выбираем случайно, то полагаем, что все эти варианты-значения равновероятны. Тогда согласно классическому определению

вероятности относительная частота варианты $\omega_i = n_i / n = P(x_i)$ - вероятность вытащить

её из генеральной совокупности.

$$\langle X \rangle = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_m \cdot p_m = M(X)$$

То есть:

$$M(X) = \bar{x}_T$$

или что то же самое:

Оценка Θ^* : Выборочная средняя

Выборочной средней \bar{x}_B называют среднее арифметическое значение признака

выборочной совокупности (также обозначается \bar{x}_n).

$$\bar{X}_B = (x_1 + \dots + x_n) \cdot \frac{1}{n}$$

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причем $n_1 + n_2 + \dots + n_k = n$, то

$$\bar{x}_B = (n_1 x_1 + n_2 x_2 + \dots + n_k x_k) / n,$$

или

$$\bar{x}_B = \left(\sum_{i=1}^k n_i x_i \right) / n,$$

Среднее абсолютное отклонение Θ (ср. арифметическое абсолютных отклонений):

$$\theta = (\sum n_i |x_i - \bar{x}_B|) / \sum n_i$$

Свойство устойчивости X_B

Свойство устойчивости: если по нескольким выборкам достаточно большого объёма

из одной и той же генеральной совокупности будут найдены выборочные средние,

то они будут приближённо равны между собой.

$$\bar{X}_{B1} \approx \bar{X}_{B2} \approx \bar{X}_{B3} \approx \dots$$

Заметим (текст только для самых смелых умов): Вынос мозга (текст только для самых смелых умов): распределенных совокупностей равны между собой, то близость выборочных средних к генеральным не зависит от отношения объема выборки к объему генеральной совокупности. Она зависит от объема выборки: чем объем выборки больше, тем меньше выборочная средняя отличается от генеральной. Например, если из одной совокупности отобран 1% объектов, а из другой совокупности отобрано 4% объектов, причем объем первой выборки оказался большим, чем второй, то первая выборочная средняя будет меньше отличаться от соответствующей генеральной средней, чем вторая.

Требования к оценке $\Theta^* = X_B$

1. \bar{x}_B - состоятельная

по теореме Чебышева:

$$P(|\bar{X}_n - M(X)| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0$$

поэтому:

$$P(|\bar{X}_B - \bar{X}_T| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0$$

2. \bar{x}_B - несмещённая

мат. ожидание оценки:

$$M(\bar{x}_B) = M\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{\sum_{i=1}^n M(X_i)}{n} = \frac{n \cdot M(X)}{n} = M(X) = \Theta_T$$

3. \bar{x}_B - эффективная? Будет ли её дисперсия являться минимальной? Не всегда! Дисперсия оценки:

$$D(\bar{x}_B) = D\left(\sum_{i=1}^n X_i / n\right) = \frac{\sum_{i=1}^n D(X_i)}{n^2} = \frac{n \cdot D(X)}{n^2} = \frac{D(X)}{n}$$

В случае нормального распределения это действительно минимум по сравнению с другими оценками Θ

Генеральная дисперсия

Генеральной дисперсией D_G называют среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения \bar{x}_G .

Если все значения x_1, x_2, \dots, x_N признака генеральной совокупности объема N различны, то

$$D_G(X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X}_G)^2$$

При этом теоретическая дисперсия признака (случайной величины) совпадает с генеральной дисперсией:

$$D(X) = D_G(X)$$

Генеральным средним квадратическим отклонением (стандартом) называют квадратный корень из генеральной дисперсии:

$$\sigma_G = \sqrt{D_G}.$$

Оценка Θ^* : Выборочная дисперсия

Выборочной дисперсией D_B называют среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения \bar{x}_B .

Если все значения x_1, x_2, \dots, x_n признака выборки объема n различны, то

$$D_B(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X}_B)^2$$

Выборочным средним квадратическим отклонением (стандартом) называют квадратный корень из выборочной дисперсии:

$$\sigma_B = \sqrt{D_B}.$$

Коэффициент вариации V – выраженное в процентах отношение выборочного среднего квадратического отклонения к выборочной средней:

$$V = \sigma_B / \bar{x}_B \cdot 100\%$$

Требования к оценке $\Theta^* = D_B$

1. D_B - состоятельная
по известной формуле:

$$D_B = \left(\frac{1}{n} \sum_{i=1}^n X_i^2 \right) - (\bar{X}_B)^2 \xrightarrow{n \rightarrow \infty} M(X^2) - M(X)^2 = D(X) = \Theta_T$$

2. D_B - смещённая!
мат. ожидание оценки:

$$M[D_B(X)] = \frac{n-1}{n} \cdot D(X)$$

т.е. при использовании этой оценки будет возникать систематическая ошибка в меньшую сторону! потому что $M(D_B) < D(X)$.

3. D_B - эффективная? Будет ли её дисперсия являться минимальной? Нет.
Даже в случае нормального распределения.
Однако при $n \rightarrow \infty$ её мат. ожидание асимптотически приближается к $D(X)$,
поэтому выборочная дисперсия является **асимптотически эффективной**.

Вывод мат. ожидания D_B

Пересчитаем выборочную дисперсию, сделав центрированную величину $X - M(X)$:

$$\begin{aligned} D_B(X) &= \overline{X^2} - (\bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (x_i - M(X) + M(X) - \bar{X})^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2 - \frac{2}{n} \sum_{i=1}^n (x_i - M(X))(M(X) - \bar{X}) + \frac{1}{n} \sum_{i=1}^n (M(X) - \bar{X})^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2 + \frac{2}{n} (M(X) - \bar{X})(n\bar{X} - nM(X)) + \frac{1}{n} n \cdot (M(X) - \bar{X})^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2 - 2(M(X) - \bar{X})^2 + (M(X) - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2 - (M(X) - \bar{X})^2 \end{aligned}$$

Теперь посчитаем мат. ожидание этой дисперсии:

$$M[D_B(X)] = M\left[\frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2 - (M(X) - \bar{X})^2\right] = M\left[\frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2\right] - M[(M(X) - \bar{X})^2] =$$

используем уже доказанное: $M(\bar{X}) = M(X)$

$$= \frac{1}{n} \sum_{i=1}^n M[(x_i - M(X))^2] - M[(M(\bar{X}) - \bar{X})^2] = \frac{1}{n} \sum_{i=1}^n D(X) - D(\bar{X}) = D(X) - \frac{D(X)}{n} = \frac{n-1}{n} D(X)$$

страшно?

Оценка генеральной дисперсии по исправленной выборочной

Зная связь: $M[D_B(X)] = \frac{n-1}{n} \cdot D(X)$

легко получить **исправленную выборочную дисперсию**: $\hat{D}_B = \frac{n}{n-1} D_B$
которая будет несмещённой точечной оценкой:

$$M[\hat{D}_B(X)] = D(X) = D_G \qquad M[D_B(X)] = \frac{n-1}{n} \cdot D(X)$$

Т.е. разница между выборочной дисперсией и исправленной лишь в множителе:

$$D_B(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X}_B)^2 \qquad \hat{D} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_B)^2$$

На практике используют исправленную дисперсию, **если** $n < 30$.

Задача 1

По результатам наблюдений случайной величины X получились числа:

1, 7, 7, 2, 3, 2, 5, 5, 4, 6, 3, 4, 3, 5, 6, 6, 5, 5, 4, 4

построить:

1. дискретный вариационный ряд
2. многоугольник частот
3. график выборочной функции распределения
4. найти выборочное среднее
5. найти выборочную дисперсию

Решение задачи 1

1. Составляем ряд распределения:

Количество элементов выборки равно 20, следовательно объем выборки $n = 20$.

Составляем ранжированный ряд:

1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 7, 7.

Выделяем варианты и их частоты $x_i(n_i)$:

1(1), 2(2), 3(3), 4(4), 5(5), 6(3), 7(2).

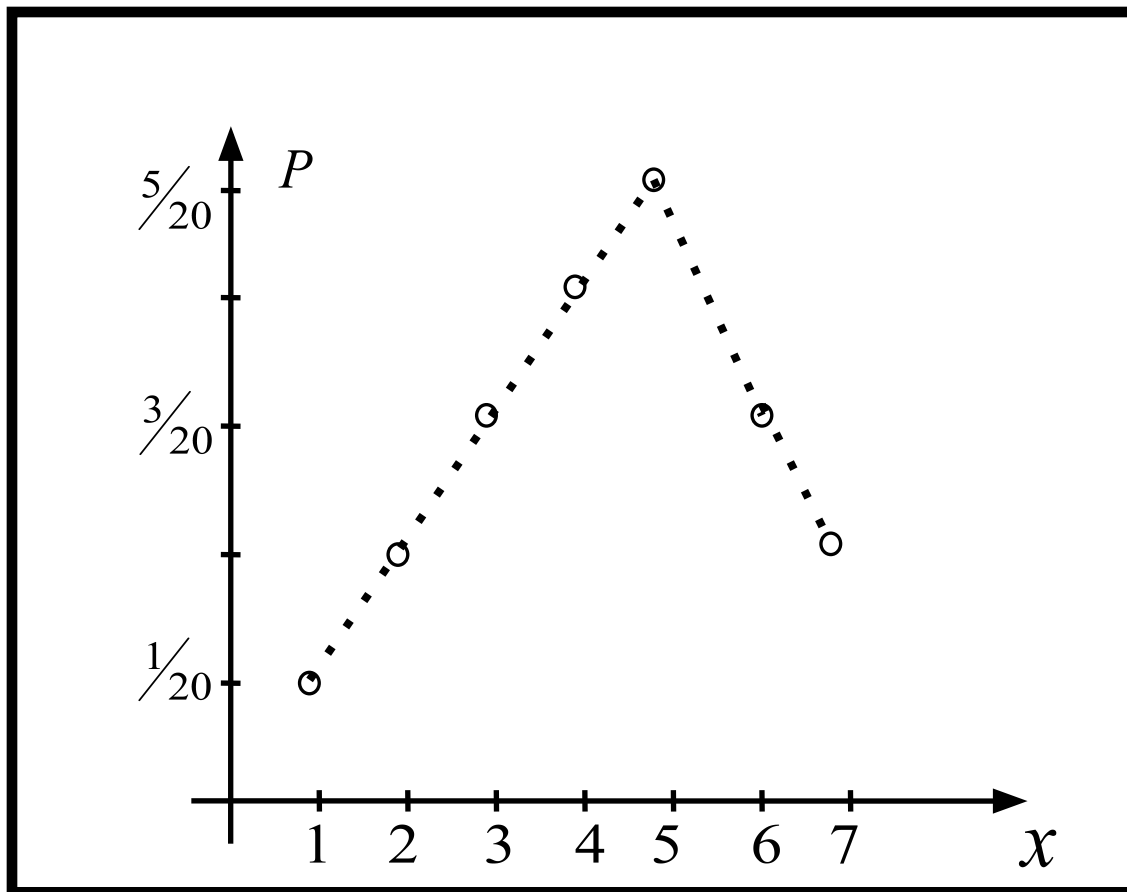
Вариантов всего 7, поэтому будем строить дискретный ряд.

Находим частоты по формуле: $p_i^* = \frac{n_i}{n}$

| | | | | | | |
|------|------|------|------|------|------|------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1/20 | 2/20 | 3/20 | 4/20 | 5/20 | 3/20 | 2/20 |

Решение задачи 1

2. По найденному ряду строим многоугольник частот:



Решение задачи 1

3. Находим выборочную функцию распределения:

$$E_*(x^1) = 0 \quad E_*(x^5) = b_*^1 = \frac{50}{1} \quad E_*(x^3) = b_*^1 + b_*^5 = \frac{50}{1} + \frac{50}{5} = \frac{50}{3}$$

$$E_*(x^4) = b_*^1 + b_*^5 + b_*^3 = \frac{50}{1} + \frac{50}{5} + \frac{50}{3} = \frac{50}{e} \quad E_*(x^2) = b_*^1 + b_*^5 + b_*^3 + b_*^4 = \frac{50}{1} + \frac{50}{5} + \frac{50}{3} + \frac{50}{4} = \frac{50}{10}$$

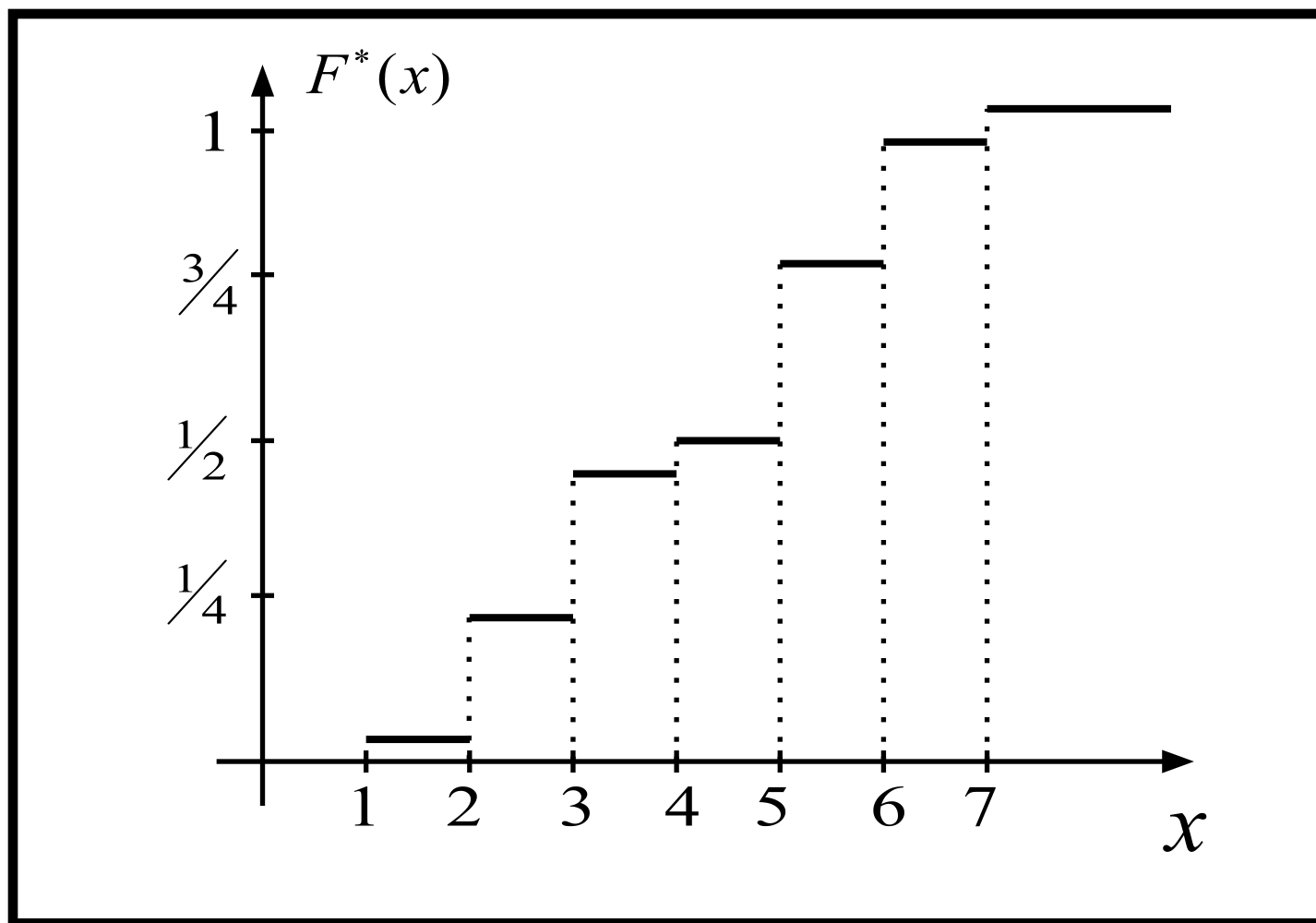
$$F^*(x_6) = p_1^* + p_2^* + p_3^* + p_4^* + p_5^* = \\ = \frac{1}{20} + \frac{2}{20} + \frac{3}{20} + \frac{4}{20} + \frac{5}{20} = \frac{15}{20}$$

$$F^*(x_7) = p_1^* + p_2^* + p_3^* + p_4^* + p_5^* + p_6^* = \\ = \frac{1}{20} + \frac{2}{20} + \frac{3}{20} + \frac{4}{20} + \frac{5}{20} + \frac{3}{20} = \frac{18}{20}$$

$$F^*(x_8) = p_1^* + p_2^* + p_3^* + p_4^* + p_5^* + p_6^* + p_7^* = \\ = \frac{1}{20} + \frac{2}{20} + \frac{3}{20} + \frac{4}{20} + \frac{5}{20} + \frac{3}{20} + \frac{2}{20} = \frac{20}{20} = 1$$

Решение задачи 1

...строим график выборочной функции распределения:



Решение задачи 1

4. Находим выборочное среднее по ряду распределения:

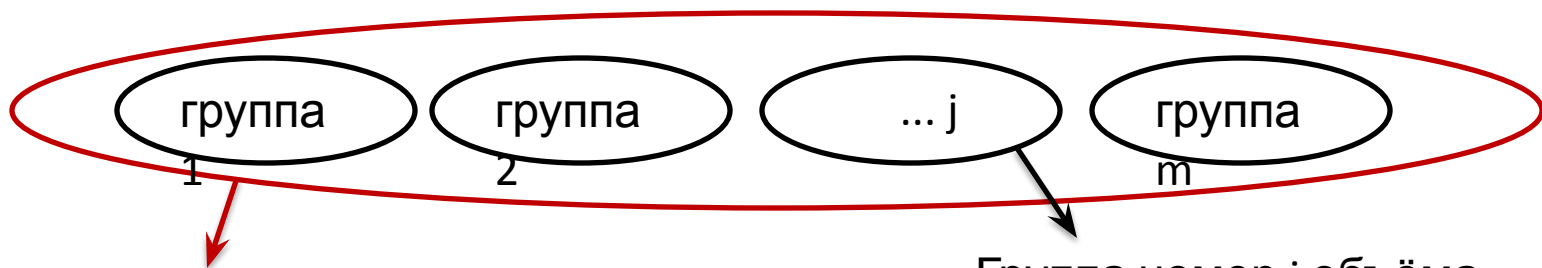
$$\bar{X}_B = 1 \cdot \frac{1}{20} + 2 \cdot \frac{2}{20} + 3 \cdot \frac{3}{20} + 4 \cdot \frac{4}{20} + 5 \cdot \frac{5}{20} + 6 \cdot \frac{3}{20} + 7 \cdot \frac{2}{20} = 4.35$$

5. Находим выборочную дисперсию:

$$D_B = 1^2 \cdot \frac{1}{20} + 2^2 \cdot \frac{2}{20} + 3^2 \cdot \frac{3}{20} + 4^2 \cdot \frac{4}{20} + \\ + 5^2 \cdot \frac{5}{20} + 6^2 \cdot \frac{3}{20} + 7^2 \cdot \frac{2}{20} - (4.35)^2 = 4.74?$$

Разбиение статистической совокупности на группы

Допустим, что все значения количественного признака X совокупности, безразлично генеральной или выборочной, разбиты на несколько групп.



Генеральная совокупность объёма n :
общее среднее \bar{x} , общая дисперсия

$$D_{\text{общ}} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Группа номер j объёма N_j :
групповая средняя \bar{x}_j , групповая дисперсия

$$D_j = \frac{1}{N_j} \sum_{i=1}^{N_j} (x_i - \bar{x}_j)^2$$

Внутригрупповая дисперсия:
(среднее групповых дисперсий)

$$D_{\text{внутр}} = \sum_{j=1}^m \frac{N_j}{n} D_j$$

$$D_{\text{межгр}} = \sum_{j=1}^m \frac{N_j}{n} (\bar{x}_j - \bar{x})^2$$

$$D_{\text{общ}} = D_{\text{внутр}} + D_{\text{межгр}}$$

Межгрупповая дисперсия:

Задача 2

Найти внутригрупповую, межгрупповую и общую дисперсии совокупности, состоящей из трёх групп:

первая группа:

| | | | |
|-------|----|----|---|
| X_i | 1 | 2 | 8 |
| n_i | 30 | 15 | 5 |

вторая группа:

| | | |
|-------|----|----|
| X_i | 1 | 6 |
| n_i | 10 | 15 |

третья группа:

| | | |
|-------|----|---|
| X_i | 3 | 8 |
| n_i | 20 | 5 |

Отв. $D_{внгр} = 4,6$; $D_{межгр} = 1$; $D_{общ} = 5,6$.

Оценка корреляции величин

Пусть данные наблюдений за признаками X и Y сведены в корреляционную таблицу.

Можно считать, что наблюдаемые Y разбиты на группы, соответствующие отдельным X.

Тогда условные средние $\langle y_x \rangle$ можно назвать групповыми средними.

Общая дисперсия в этом случае может быть представлена в виде суммы дисперсий:

| Y | X | |
|-------------|-----|-----|
| | 3 | 9 |
| 3 | 4 | 13 |
| 5 | 6 | 7 |
| n_x | 10 | 20 |
| \bar{y}_x | 4,2 | 3,7 |

$$D_{\text{общ}} = D_{\text{внутр}} + D_{\text{межгр}}$$

Можно доказать, что:

1. если Y связан с X функциональной зависимостью, то $D_{\text{межгр}} / D_{\text{общ}} = 1$
2. если Y связан с X корреляционной зависимостью, то $D_{\text{межгр}} / D_{\text{общ}} < 1$

(функциональная зависимость является крайним случаем вероятностной).

$$\eta_{yx} = \sqrt{D_{\text{межгр}} / D_{\text{общ}}}$$

– выборочное корреляционное отношение Y к X.

Ассиметрия и Эксцесс

Начальный момент порядка k случайной величины X $\nu_k = M(X^k)$
это математическое ожидание величины X^k :

Центральный момент порядка k случайной величины X $\mu_k = M[(X - M(X))^k]$
это математическое ожидание величины $(X - M(X))^k$:

Начальные и центральные эмпирические моменты являются состоятельными оценками соответственно начальных и центральных теоретических моментов.

Асимметрия эмпирического распределения: $a_s = \mu_3 / \sigma^3$
 $a_s > 0$ (< 0), если правый хвост распределения длиннее (короче) левого.

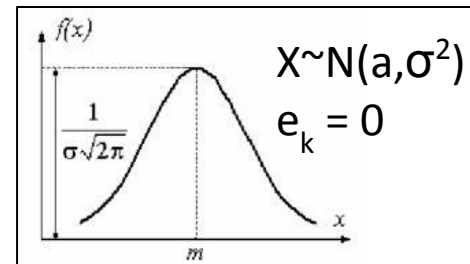
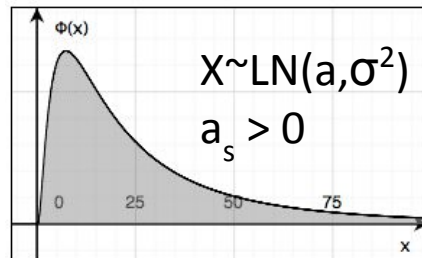
Эксцесс эмпирического распределения: $e_k = \mu_4 / \sigma^4 - 3$
 $e_k > 0$ (< 0), если пик возле мат. ожидания острый (гладкий).

Логнормальное
распределение

:

$X \sim \text{LN}(a, \sigma^2) \Rightarrow$

$\Rightarrow \ln(X) \sim N(a, \sigma^2)$



Метод моментов для точечной оценки параметров распределения

(предложен
Пирсоном)

А. Оценка одного параметра. Пусть задан вид плотности распределения $f(x, \theta)$, определяемой одним неизвестным параметром θ . Требуется найти точечную оценку параметра θ .

Для оценки одного параметра достаточно иметь одно уравнение относительно этого параметра. Следуя методу моментов, приравняем, например, начальный теор-

$$M(X) = \bar{x}_n.$$

Математическое ожидание $M(X)$, как видно из соотношения

$$M(X) = \int_{-\infty}^{\infty} xf(x; \theta) dx = \varphi(\theta),$$

есть функция от θ , поэтому (*) можно рассматривать как уравнение с одним неизвестным θ . Решив это уравнение относительно параметра θ , тем самым найдем его точечную оценку θ^* , которая является функцией от выборочной средней, следовательно, и от вариантов выборки:

$$\theta^* = \psi(x_1, x_2, \dots, x_n).$$

Задача

Пример 1. Найти методом моментов по выборке x_1, x_2, \dots, x_n точечную оценку неизвестного параметра λ показательного распределения, плотность распределения которого $f(x) = \lambda e^{-\lambda x}$ ($x \geq 0$).

Решени

е:

$$1/\lambda = \bar{x}_B.$$

Б. Оценка двух параметров методом Пирсона

Б. Оценка двух параметров. Пусть задан вид плотности распределения $f(x; \theta_1, \theta_2)$, определяемой неизвестными параметрами θ_1 и θ_2 . Для отыскания двух параметров необходимы два уравнения относительно этих параметров. Следуя методу моментов, приравняем, например, начальный теоретический момент первого порядка начальному эмпирическому моменту первого порядка и центральный теоретический момент второго порядка центральному эмпирическому моменту второго порядка:

$$\left. \begin{aligned} M(X) &= \bar{x}_B, \\ D(X) &= D_B. \end{aligned} \right\}$$

Математическое ожидание и дисперсия есть функции от θ_1 и θ_2 , поэтому (***) можно рассматривать как систему двух уравнений с двумя неизвестными θ_1 и θ_2 . Решив

Задача

Пример 2. Найти методом моментов по выборке x_1, x_2, \dots, x_n точечные оценки неизвестных параметров a и σ нормального распределения

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

Решени

е.
Приняв во внимание, что математическое ожидание нормального распределения равно параметру a , дисперсия равна σ^2 (см. гл. XII, § 2), имеем:

$$a = \bar{x}_B, \quad \sigma^2 = D_B.$$

З а м е ч а н и е 1. Для оценок неизвестных параметров можно приравнивать не только сами моменты, но и функции от моментов. В частности, этим путем получают состоятельные оценки характеристик распределений, которые являются функциями теоретических моментов. Например, асимметрия теоретического распределения (см. гл. XII, § 9)

$$A_S = \mu_3 / \sigma^3 = \mu_3 / (\sqrt{\mu_2})^3$$

есть функция от центральных моментов второго и третьего порядков. Заменяя эти теоретические моменты соответствующими эмпирическими моментами, получим точечную оценку асимметрии

$$A_S^* = m_3 / (\sqrt{m_2})^3.$$

Метод наибольшего правдоподобия

(предложен
Фишером)

А. Дискретные случайные величины. Пусть X — дискретная случайная величина, которая в результате n испытаний приняла значения x_1, x_2, \dots, x_n . Допустим, что вид закона распределения величины X задан, но неизвестен параметр θ , которым определяется этот закон. Требуется найти его точечную оценку.

Обозначим вероятность того, что в результате испытания величина X примет значение x_i ($i = 1, 2, \dots, n$), через $p(x_i; \theta)$.

Функцией правдоподобия дискретной случайной величины X называют функцию аргумента θ :

$$L(x_1, x_2, \dots, x_n; \theta) = p(x_1; \theta) p(x_2; \theta) \dots p(x_n; \theta),$$

где x_1, x_2, \dots, x_n — фиксированные числа.

L - это вероятность получить такую выборку!

В качестве точечной оценки параметра θ принимают такое его значение $\theta^* = \theta^*(x_1, x_2, \dots, x_n)$, при котором функция правдоподобия достигает максимума. Оценку θ^* называют *оценкой наибольшего правдоподобия*.

Метод наибольшего правдоподобия

Логарифмической функцией правдоподобия называют функцию $\ln L$. Как известно, точку максимума функции $\ln L$ аргумента θ можно искать, например, так:

1) найти производную $\frac{d \ln L}{d\theta}$;

2) приравнять производную нулю и найти критическую точку — корень полученного уравнения (его называют *уравнением правдоподобия*);

3) найти вторую производную $\frac{d^2 \ln L}{d\theta^2}$; если вторая производная при $\theta = \theta^*$ отрицательна, то θ^* — точка максимума.

Недостаток метода состоит в том, что он часто требует сложных вычислений.

З а м е ч а н и е 2. Оценка наибольшего правдоподобия не всегда совпадает с оценкой, найденной методом моментов.

Метод наибольшего правдоподобия

Б. Непрерывные случайные величины.

Функцией правдоподобия непрерывной случайной величины X называют функцию аргумента θ :

$$L(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta),$$

где f - плотности распределений вероятности величин x .

Задача

Пример 3. Найти методом наибольшего правдоподобия оценку параметра λ показательного распределения

$$f(x) = \lambda e^{-\lambda x} \quad (0 < x < \infty),$$

если в результате n испытаний случайная величина X , распределенная по показательному закону, приняла значения x_1, x_2, \dots, x_n .

Решение

Решение. Составим функцию правдоподобия, учитывая, что $\theta = \lambda$:

$$L = f(x_1; \lambda) f(x_2; \lambda) \dots f(x_n; \lambda) = (\lambda e^{-\lambda x_1}) (\lambda e^{-\lambda x_2}) \dots (\lambda e^{-\lambda x_n}).$$

Отсюда

$$L = \lambda^n e^{-\lambda \sum x_i}.$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L = n \ln \lambda - \lambda \sum x_i.$$

Найдем первую производную по λ :

$$\frac{d \ln L}{d\lambda} = \frac{n}{\lambda} - \sum x_i.$$

Напишем уравнение правдоподобия, для чего приравняем первую производную нулю:

$$(n/\lambda) - \sum x_i = 0.$$

Найдем критическую точку, для чего решим полученное уравнение относительно λ :

$$\lambda = n / \sum x_i = 1 / (\sum x_i / n) = 1 / \bar{x}_B.$$

Найдем вторую производную по λ :

$$\frac{d^2 \ln L}{d\lambda^2} = -\frac{n}{\lambda^2}.$$

Легко видеть, что при $\lambda = 1/\bar{x}_B$ вторая производная отрицательна;

Откуда следует, $\lambda^* = 1/\bar{x}_B$.

что:

З а м е ч а н и е. Если плотность распределения $f(x)$ непрерывной случайной величины X определяется двумя неизвестными параметрами θ_1 и θ_2 , то функция правдоподобия является функцией двух независимых аргументов θ_1 и θ_2 :

$$L = f(x_1; \theta_1, \theta_2) f(x_2; \theta_1, \theta_2) \dots f(x_n; \theta_1, \theta_2),$$

где x_1, x_2, \dots, x_n — наблюдавшиеся значения X . Далее находят логарифмическую функцию правдоподобия и для отыскания ее максимума составляют и решают систему

$$\begin{cases} \frac{\partial \ln L}{\partial \theta_1} = 0, \\ \frac{\partial \ln L}{\partial \theta_2} = 0. \end{cases}$$

Пример 4. Найти методом наибольшего правдоподобия оценки параметров a и σ нормального распределения

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2},$$

если в результате n испытаний величина X приняла значения x_1, x_2, \dots, x_n .

Решение. Составим функцию правдоподобия, учитывая, что $\theta_1 = a$ и $\theta_2 = \sigma$:

$$L = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_1-a)^2/2\sigma^2} \cdot \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_2-a)^2/2\sigma^2} \dots \times \\ \dots \times \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_n-a)^2/2\sigma^2}$$

Отсюда

$$L = \frac{1}{\sigma^n (\sqrt{2\pi})^n} e^{-\sum (x_i - a)^2 / 2\sigma^2}$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L = -n \ln \sigma + \ln \frac{1}{(\sqrt{2\pi})^n} - \frac{\sum (x_i - a)^2}{2\sigma^2}.$$

Найдем частные производные по a и по σ :

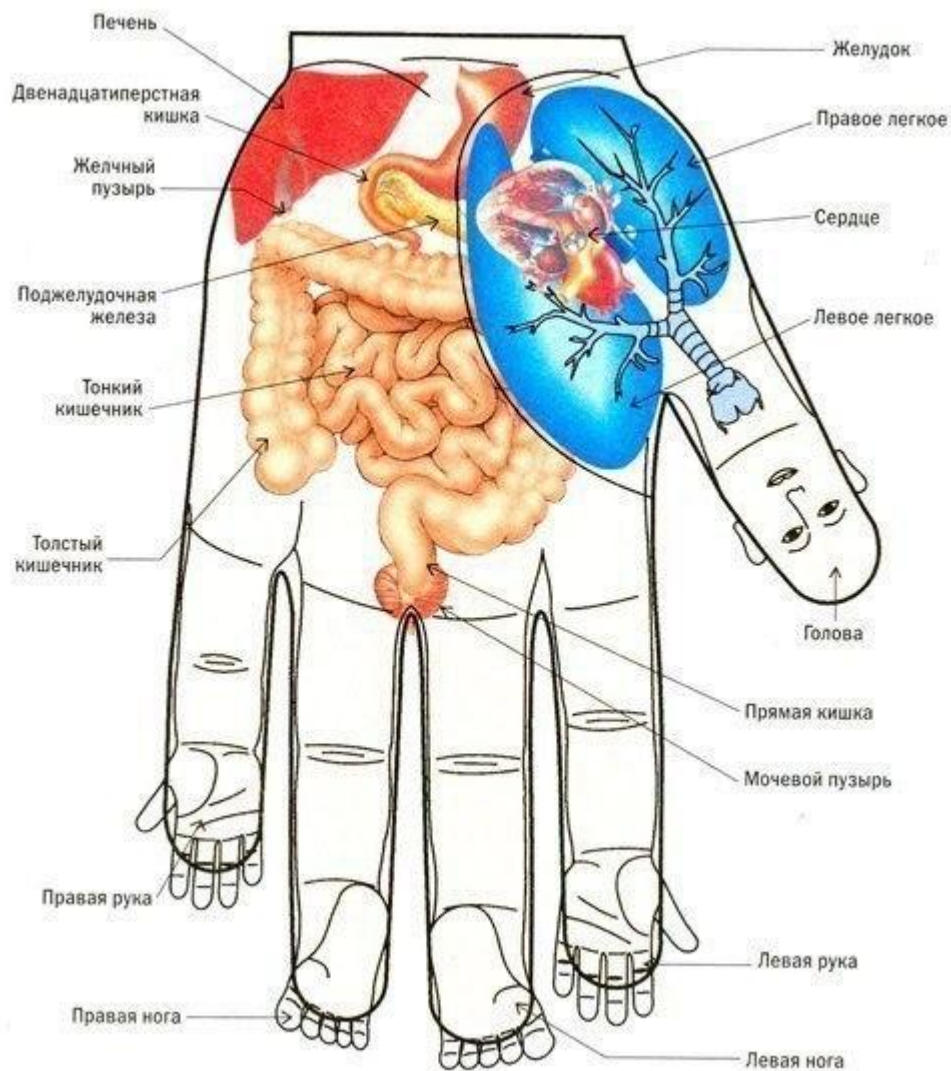
$$\frac{\partial \ln L}{\partial a} = \frac{\sum x_i - na}{\sigma^2}; \quad \frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum (x_i - a)^2}{\sigma^3}.$$

Приравняв частные производные нулю и решив полученную систему двух уравнений относительно a и σ^2 , получим:

$$a = \sum x_i / n = \bar{x}_B; \quad \sigma^2 = (\sum (x_i - \bar{x}_B)^2) / n = D_B.$$

Итак, искомые оценки наибольшего правдоподобия: $a^* = \bar{x}_B$; $\sigma^* = \sqrt{D_B}$. Заметим, что первая оценка несмещенная, а вторая смещенная.

Спасибо за внимание :)



...точечные
оценки

Вопросы для контроля усвояемости предмета (стр1)

- Необходимость статистической оценки теоретического распределения
- Статистическая оценка
- Точечная статистическая оценка
- Точечная оценка Θ – случайная величина?
- Состоятельная статистическая оценка
- Несмещённая статистическая оценка
- Эффективная статистическая оценка
- Как связана генеральная средняя с мат. ожиданием признака?
- Выборочная средняя
- Свойство устойчивости выборочной средней
- Среднее абсолютное отклонение
- Является ли выборочная средняя состоятельной?
- Является ли выборочная средняя несмещённой?
- Является ли выборочная средняя эффективной?
- Генеральная дисперсия
- Выборочная дисперсия
- Генеральное и выборочное среднее квадратическое отклонение

Вопросы для контроля усвояемости предмета (стр2)

- Коэффициент вариации
- Является ли выборочная дисперсия состоятельной?
- Является ли выборочная дисперсия несмещённой?
- Является ли выборочная дисперсия эффективной?
- Исправленная выборочная дисперсия
- Групповая средняя, групповая дисперсия
- Внутригрупповая дисперсия
- Межгрупповая дисперсия
- Связь внутригрупповой и межгрупповой дисперсий
- Начальный момент порядка k
- Центральный момент порядка k
- Асимметрия
- Эксцесс
- Метод моментов. Оценка одного или двух параметров
- Метод наибольшего правдоподобия
- Функция правдоподобия