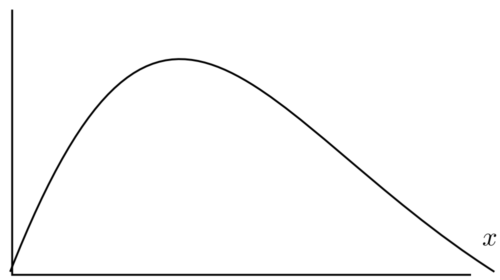
1 Занятие 1

1.1 Основные распределения в Мат Стат

1.1.1 Гамма распределение

$$\xi \sim \Gamma(\lambda, a)\ \lambda > 0, a > 0$$
$$P(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}\ \{(0, +\infty)\}$$



$$\Gamma(S+1) = S\Gamma(S) \qquad \Gamma(S) = \int_0^\infty x^{s-1} e^{-x} dx\ x > 0$$
$$M[\Gamma] = \int_{-\infty}^{\infty} \rho(x) dx = \int_0^\infty \frac{\lambda^a}{\Gamma(a)} (\frac{t}{\lambda})^a e^{-t} \frac{dt}{\lambda} == \frac{1}{\lambda\Gamma(a)} \int_0^\infty t^a e^{-t} dt = \frac{a}{\lambda}$$
$$D[\xi] = M[\xi^2] - M^2[\xi] = \frac{a^2 + a}{\lambda^2} - (\frac{a}{\lambda})^2 = \frac{a}{\lambda^2}$$

Теорема 1.1 (Свойство суммы). $\xi_1, \dots, \xi_n$ независимы, $\xi_i \sim \Gamma(\lambda, a_i)$, $\eta = \xi_1 + \dots + \xi_n \sim \Gamma(\lambda, a_1 + \dots + a_n)$

Доказательство. $\xi_1 \sim \Gamma(\lambda, a_1)$. $\xi_2 \sim \Gamma(\lambda, a_2)$ - независимые, $\eta = \xi_1 + \xi_2$

$$\Phi(y) = P(\eta < y) = P(\xi_2 + \xi_2 < y) = \iint_{x_1+x_2<y} p(x_1, x_2) dx_1 dx_2 =$$
$$= \int_0^y dx_2 \int_0^{y-x_1} \frac{\lambda^{a_1}}{\Gamma(a_1) x_1^{a_1-1} e^{-\lambda x_1}} \frac{\lambda^{a_2}}{\Gamma(a_2) x_2^{a_2-2} e^{-\lambda x_2}} dx_2$$
$$\varphi(y) = \Phi'(y)$$

□

1.1.2 Распределение Парсона χ^2

$\xi_i \sim N(0,1)$ - независимы, $\eta = \xi_1^2 + \cdots + \xi_n^2 = \chi^2$

$$\Phi(y) = P(\xi^2 < y) = \begin{cases} y \le 0 & : 0 \\ y > 0 & : P(-\sqrt{y} < \xi < \sqrt{y}) \end{cases}$$
$$\varphi(x) = \begin{cases} \frac{1}{2\sqrt{y}} F'(\sqrt{y}) + \frac{1}{2\sqrt{y}} F'(-\sqrt{y}),\ y > 0 \\ 0,\ y < 0 \end{cases}$$
$$p(x) = \frac{e^{x^2/2}}{\sqrt{2\pi}}$$
$$\varphi(y) = \begin{cases} \frac{1}{\sqrt{y}} \frac{e^{-y/2}}{\sqrt{2\pi}},\ y > 0 \\ 0,\ y \le 0 \end{cases}$$
$$p(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \{(0; +\infty)\} \qquad \lambda = \frac{1}{2} \qquad a = \frac{1}{2}$$
$$\xi^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2}) \qquad \xi_1^2 + \cdots + \xi_n^2 \sim \Gamma(\frac{1}{2}, \frac{n}{2}) = \chi^2(n)$$

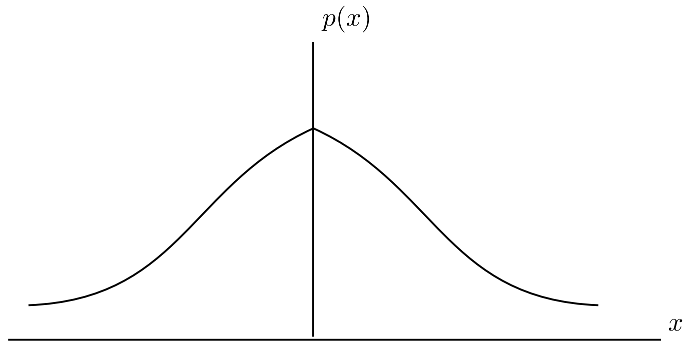
n - число степеней свободы

$$M[\eta] = \frac{a}{\lambda} = \frac{n/2}{1/2} = n$$
$$D[\eta] = \frac{a}{\lambda^2} = \frac{n/2}{1/4} = 2n$$

Теорема 1.2 (Свойство суммы)**.** $\xi_1, \ldots, \xi_m$ - независ, $\xi_i \sim \chi^2(n_i)$, $\xi_1 + \cdots + \xi_n \sim \chi^2(n_1 + \cdots + n_m)$

1.2 Распределение Стьюдента (Госсет)

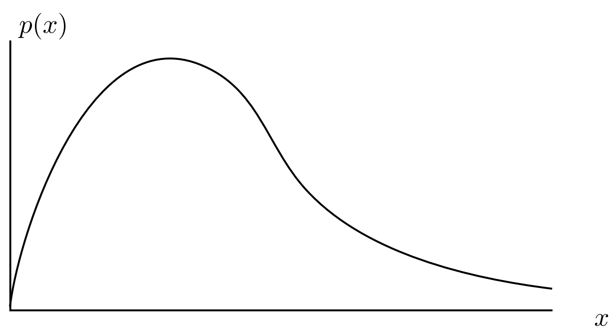
$\xi \sim N(0,1)$, $\eta \sim \chi^2(m)$ - независимы, $\frac{\xi}{\sqrt{\eta/m}} \sim t(m)$



$$p(x) = \frac{(m)^{m/2}\Gamma(\frac{m+1}{2})}{\sqrt{\pi}\Gamma(\frac{m}{2})(x^2+m)^{\frac{m+1}{2}}}$$

1.3 Распределение Фишера

$\xi \sim \chi^2(n)$, $\eta \sim \chi^2(m)$ - независимые, $\frac{\xi/n}{\eta/m} \sim F(n,m)$



1.4 Нормальное распределение

$$p(\vec{x}) = \frac{1}{(\sqrt{2\pi})^n}\frac{1}{\sqrt{detK}}e^{-\frac{1}{2}(\vec{x}-\vec{a})^T K^{-1}(\vec{x}-\vec{a})}$$
$$\vec{\xi} \sim N(\vec{a}, R)$$

Свойства:

- $\xi \sim N(0,1)$, $\eta = a\xi + b \sim N(b, a^2)$
- $\xi \sim N(\alpha, \sigma^2)$, $\eta = a\xi + b \sim N(a\alpha + b, \sigma^2 a^2)$
- $\xi \sim N(\vec{0}, E)$,$\vec{\eta} = A\vec{\xi} + \vec{b}$, $A : n \times n, detA \neq 0$

$$\Phi(t_1, \dots, t_n) = P(\eta_1 < t_1, \dots, \eta_n < t_n) = P(\vec{\eta} < \vec{t}) = P(A\vec{\xi} + \vec{b} < \vec{t}) =$$
$$= \int \cdots \int_{A\vec{x}+\vec{b}<\vec{t}} p(x_1, \dots, x_n) dx_1 \dots dx_n =$$
$$\vec{y} = A\vec{x} + \vec{b} \qquad J = \left|\frac{\partial \vec{x}}{\partial \vec{y}}\right| \qquad \frac{1}{J} = detA$$
$$= \int \cdots \int_{\vec{y}<\vec{t}} p(A^{-1}(\vec{y} - \vec{b}))\frac{1}{|detA|} dy_1 \dots dy_n$$
$$\varphi(\vec{t}) = p(A^{-1}(\vec{y} - \vec{b}))\frac{1}{|detA|}$$
$$\varphi(\vec{t}) = \frac{1}{|detA|}\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}(A^{-1}(\vec{y}-\vec{b})^T)(A^{-1}(\vec{y}-\vec{b}))} =$$
$$= \frac{1}{|detA|}\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}(\vec{t}-\vec{b})^T (A^T)^{-1} A^{-1}(\vec{t}-\vec{b})}$$
$$K = AA^T \qquad \vec{\eta} = A\vec{\xi} + \vec{b} \sim N(\vec{b}, AA^T)$$

- $\xi \sim N(\vec{a}, K)$, $\vec{\eta} = A\vec{\xi} + \vec{b} \sim N(A\vec{a} + \vec{b}, AKA^T)$, $A : n \times n$, $detA \neq 0$
- Для $A : m \times n$ два предыдущих свойства так же верны
- ξ, η - независ $\Rightarrow cov(\xi, \eta) = 0$, в другую сторону не верно

$$\begin{cases} \xi \sim N(a_1, \sigma_1^2) \\ \eta \sim N(a_2, \sigma_2^2) \\ \text{независимые} \end{cases} \Leftrightarrow (\xi, \eta) \sim N\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}\right)$$

$$\begin{cases} \xi \sim N(a_1, \sigma_1^2) \\ \eta \sim N(a_2, \sigma_2^2) \\ cov(\xi, \eta) = 0 \end{cases} \Leftarrow (\xi, \eta) \sim N\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}\right)$$

Лемма 1.1 (Лемма Фишера)**.** Пусть $\vec{\xi} \sim N(\vec{0}, E)$ и C ортогональная матрица, $\vec{\eta} = C\vec{\xi}$, тогда $\forall k = 1 \dots n-1$ сл. вел. $\varkappa = \sum_{i=1}^n \xi_i^2 - \eta_1^2 - \eta_2^2 - \dots - \eta_k^2 \sim \chi^2(n-k)$ и вел $\varkappa, \eta_1, \eta_2, \dots, \eta_k$ независ.

Доказательство.

$$\vec{\eta} \sim N(\vec{0}, \underbrace{CC^T}_{E})$$

$$\eta_1^2 + \dots + \eta_n^2 = \vec{\eta}^T\vec{\eta} = \vec{\xi}^T C^T C \vec{\xi} = \xi_1^2 + \dots + \xi_n^2$$

$$\varkappa = \eta_{k-1}^2 + \dots + \eta_n^2$$

$$\varkappa = \chi^2(n-k)$$

□

Теорема 1.3 (Фишера)**.** Пусть $\xi_1, \dots, \xi_n$ независ и $\xi_i \sim N(a, \sigma^2)$, тогда:

1. $\varphi = \sqrt{n}\frac{\bar{\xi}-a}{\sigma} \sim N(0,1)$, $\bar{\xi} = \frac{1}{n}\sum_1^n \xi_i$
2. $\psi = \sum_{i=1}^n \frac{(\xi_i - \bar{\xi})^2}{\sigma^2} \sim \chi^2(n-1)$
3. φ и ψ независ.

Доказательство.

$$\varphi = \frac{1}{\sqrt{n}} \frac{\sum^n \xi_i - na}{\sigma} = \frac{1}{\sqrt{n}} \sum^n \left(\frac{\xi_i - a}{\sigma} \right)$$

$$\frac{\xi_i - a}{\sigma} = \frac{1}{\sigma}\xi_i - \frac{a}{\sigma} \sim N(\frac{a}{\sigma} - \frac{a}{\sigma}, \sigma^2 \frac{1}{\sigma^2}) = N(0,1)$$

$$\varphi = \frac{1}{\sqrt{n}} \sum^n \eta_i = (\frac{1}{\sqrt{n}} \dots \frac{1}{\sqrt{n}})\vec{\eta} \sim N(\vec{0}, AA^T) = N(0,1)$$

1) Доказан

$$\psi = \sum^n \left(\underbrace{\frac{\xi_i - a}{\sigma}}_{\eta_i \sim N(0,1)} - \underbrace{\frac{\bar{\xi} - a}{\sigma}}_{\bar{\eta}} \right)^2 = \sum^n (\eta_i - \bar{\eta})^2 = \sum^n (\eta_i^2 - 2\eta_i\bar{\eta} + (\bar{\eta})^2) =$$

$$= \sum^n \eta_i^2 - 2\bar{\eta}\sum^n \eta_i + n(\bar{\eta})^2 = \sum^n \eta_i^2 - n(\bar{\eta})^2$$

$$\eta_i \sim N(0,1) \qquad \zeta^2 = n\bar{\eta}^2 \qquad \zeta = \sqrt{n}\bar{\eta} = \frac{1}{\sqrt{n}}\sum^n \eta_i = A\vec{\eta} = \varphi$$

$A = \left(\frac{1}{\sqrt{n}} \dots \frac{1}{\sqrt{n}} \right) \Rightarrow C$ - ортог. матрица (Грамма-Шмидта)

(A получается строчкой матрицы C и тогда ζ - одна из координат в другом базисе и применима Лемма Фишера)

По лемме Фишера $\psi \sim \chi^2(n-1)$, ψ и $A\vec{\eta}$ независ $\square$

Теорема 1.4 (О проекции). Пусть $\vec{\xi} \sim N(\vec{0}, \sigma^2 E)$, $L_1 : dim L_1 = m_1$ и $L_2 : dim L_2 = m_2$ два ортогональных подпространства $\mathbb{R}^n$, $\vec{\eta_1}$ - проекция $\vec{\xi}$ на L_1, норм. распр., независ. и $\frac{|\eta_1|^2}{\sigma^2} \sim \chi^2(dim L_1)$, $\frac{|\eta_2|^2}{\sigma^2} \sim \chi^2(dim L_2)$

Доказательство. $\vec{\eta_1} = A_1 \vec{\xi} \sim N(\dots, \dots)$, $\vec{\zeta} = C\vec{\xi}$, C - ортогональная. $\vec{\zeta} \sim N(\vec{0}, C\sigma^2 E C^T) = N(\vec{0}, \sigma^2 E)$. Новый ортонормированный базис $e'_1 \dots e'_m$ в L_1, $e'_{m+1} \dots e'_n$ в L_2, $\vec{\eta_1} = \zeta_1 e'_1 + \dots + \zeta_m e'_m$, $\vec{\eta_2} = \zeta_{m+1} e'_{m+1} + \dots + \zeta_n e'_n$

$$\frac{\bar{\xi}}{\sigma} \sim N(\vec{0}, E) \qquad \frac{|\eta_1|^2}{\sigma^2} = \sum \frac{\xi_i^2}{\sigma^2} \sim \chi^2(m_1)$$

$\square$

2 Порядковые случайные величины

$\xi_1, \xi_2, \dots, \xi_n$ назавис, $\xi_i \sim F(x)$

$$\eta = min(\xi_1, \dots \xi_n) \sim ? \qquad \zeta = max(\xi_1, \dots \xi_n) \sim ?$$

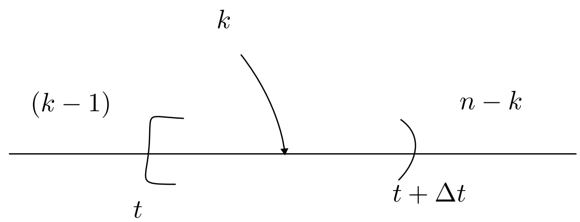
$$\Phi(y) = P(\eta < y) = 1 - P(\eta \geq y) = 1 - P(min(\xi_1, \dots, \xi_n) \geq y) =$$
$$= 1 - P(\xi_1 \geq y, \dots, \xi_n \geq y) = 1 - \prod_{i=1}^{n} P(\xi_i \geq y) =$$
$$= 1 - \prod_{i=1}^{n} (1 - P(\xi_i < y)) = 1 - (1 - F(y))^n$$

$$\Psi(z) = P(\zeta < z) = P(max(\xi_1, \dots, \xi_n) < z) =$$
$$= P(\xi_1 < z, \dots, \xi_n < z) = \prod_{i=1}^{n} P(\xi_i < z) = (F(z))^n$$

Порядковые величины:

$$\xi_{(1)} = min(\xi_1, \dots, \xi_n)$$
$$\xi_{(2)} = min(\xi_i : \xi_i \neq \xi_{(1)})$$
$$\xi_{(3)} = min(\xi_i : \xi_i \neq \xi_{(1)}, \xi_i \neq \xi_{(2)})$$
$$\dots$$
$$\xi_{(n)} = max(\xi_1, \dots, \xi_n)$$

Положим $F(x)$ - непрерывна:



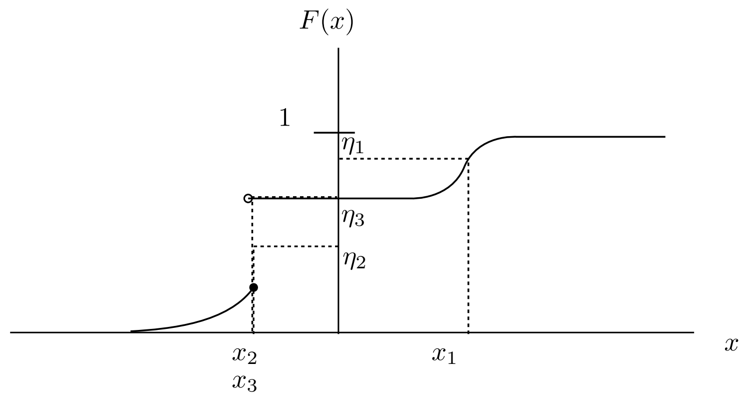
$$P(t \ \leq \xi_{(k)} < t+\Delta t) =$$
$$= nP(t \leq \xi < t+\Delta t) C_{n-1}^{k-1} (P(\xi < t))^{k-1} (P(\xi \geq t+\Delta t))^{n-k} C_{n-k}^{n-k} =$$
$$= \varkappa(t+\Delta t) - \varkappa(t)$$
$$n\frac{F(t+\Delta t)-F(t)}{\Delta t} C_{n-1}^{k-1} (F(t))^{k-1} (1-F(t+\Delta t))^{n-k} = \frac{\varkappa(t+\Delta t)-\varkappa(t)}{\Delta t}$$
$$\varkappa(t) = np(t) C_{n-1}^{k-1} (1-F(t))^{n-k} (F(t))^{k-1}$$

$\varkappa(t)$ - плотность распределения $\xi_{(k)}$, $p(t) = F'(t)$

$\xi_{(1)}$ и $\xi_{(n)}$ совместное распр

$$G(y,z) = P(\xi_{(1)} < y, \xi_{(n)} < z)$$
$$P(\xi_{(n)} < z) = P(\xi_{(1)} < y, \xi_{(n)} < z) + P(\xi_{(1)} \geq y, \xi_{(n)} < z)$$
$$\Psi(z) = (F(z))^n = P(\xi_{(1)} < y, \xi_{(n)} < z) + \prod_{i=1}^{n} \underbrace{P(y \leq \xi_i < z)}_{F(z)-F(y)}$$
$$G(y,z) = P(\xi_{(1)} < y, \xi_{(n)} < z) = \begin{cases} (F(z))^n, & y > z \\ (F(z))^n - (F(z)-F(y))^n, & y \leq z \end{cases}$$

3 Моделирование случайных величин



$\xi \sim F(x)$, $\eta \sim R(0,1)$, $F(x) = \eta_1 \to x_1$, для псевдослучайных чисел вихрь Мерсенна

4 Основные задачи статистики

Явление $\to$ математическая модель явления $\to$ вероятностная модель явления ξ_i.

Выборка - n наблюдений над явлением $\to$описательная статистика (непараметрическая), выбор классов (e.g. $\varepsilon \sim N(a, \sigma^2)$) $\to$ параметрическая статистика (e.g. $\varepsilon \sim N(a, \sigma^2)$, $a-?$, $b-?$)

Пример. Пытаемся понять как остывает чашка чая.

$$\frac{dT}{dt} = k(T - T_0) + \varepsilon$$

Вероятностная модель явления с двумя случайными величинами (погрешности измерений ε и внутри k).

Распределения полагаем нормальными и так далее.

1. Определение параметров и оценка их точности
2. Проверка статистических гипотиз

Характеристики модели θ, по выборке оценка $\bar{\theta}(\vec{x}_n)$, n - объём выборки.

Статистика - $\forall$ барелевская функция от $\vec{x}_n$ (борелевская $g : \mathbb{R} \to \mathbb{R}$ $\forall B \in \mathbb{B} \hookrightarrow g^{-1}(B) \in \mathbb{B}$).

Воспринимаем $\vec{x}_n$ с двух сторон:

1. конкретные наблюдения над явлением

2. независимые случайные величны с распределением, одинаковым со случайными величинами в вероятностной модели

4.1 Свойства оценок

Θ - множество значений θ, $\theta \in \Theta$, $\tilde{\theta}(\vec{x}_n)$ - оценка θ по выборке

1. несмещённость $\forall \theta \in \Theta \hookrightarrow M[\tilde{\theta}(\vec{x}_n)] = \theta$
2. состоятельность $\forall \theta \in \Theta \hookrightarrow \tilde{\theta} \xrightarrow{P} \theta$ (i.e. $\forall \varepsilon > 0 \; P(|\tilde{\theta} - \theta| \geq \varepsilon) \underset{n\to\infty}{\to} 0)$
3. сравнение оценок $\tilde{\theta}_1$ эффективнее $\tilde{\theta}_2$, если $\forall \theta \in \Theta \hookrightarrow D[\tilde{\theta}_1] \leq D[\tilde{\theta}_2]$ и $\exists \theta \in \Theta : \; D[\tilde{\theta}_1] < D[\tilde{\theta}_2]$

Теорема 4.1 (Достаточное условие состоятельности)**.** Если $\tilde{\theta}$ является несмещённой оценкой θ и $D[\tilde{\theta}] \underset{n\to\infty}{\to} 0$, то $\tilde{\theta}$ является состоятельной оценкой ($\forall \theta \in \Theta$)

Доказательство. $M[\tilde{\theta}] = \theta$, по неравенству Чебышёва:

$$\forall \varepsilon > 0 \hookrightarrow P(|\tilde{\theta} - \underbrace{M[\tilde{\theta}]}_{\theta}| < \varepsilon) \geq 1 - \frac{D[\tilde{\theta}]}{\varepsilon^2} \underset{n\to\infty}{\to} 1$$

$\square$

Задача (T1)**.** Пытаемся понять по двум серийным номерам сколько всего танков.

$\xi \sim R(0, \theta)$, $\theta > 0$ вер. модель., $\vec{x}_n$ - выбрка объёмом n

$$\tilde{\theta}_1 = 2\bar{x} = 2\frac{1}{n}\sum_{i=1}^{n} x_i$$
$$\tilde{\theta}_2 = \min x_i$$
$$\tilde{\theta}_3 = \max x_i$$
$$\tilde{\theta}_4 = x_1 + \frac{1}{(n-1)}\sum_{i=2}^{n} x_i$$

$$M[\xi] = \int_{-\infty}^{\infty} xp(x)dx = \int_0^\theta \frac{x}{\theta}dx = \frac{\theta}{2}$$
$$M[\xi^2] = \int_{-\infty}^{\infty} x^2p(x)dx = \int_0^\theta \frac{x^2}{\theta}dx = \frac{\theta^2}{3}$$

Рассматриваем $\tilde{\theta}_1$:

Несмещённость $\forall\theta > 0 M[\tilde{\theta}] = \theta$:

$M[\frac{\theta}{n}\sum_{i=1}^n x_i] = \frac{2}{n}\sum_{i=1}^n M[x_i] = 2M[\xi] = \theta$ несмещённая

$D[\tilde{\theta}_1] = D[\frac{2}{n}\sum x_i] = \frac{1}{n^2}\sum D[x_i] = \frac{4}{n}D[\xi] = \frac{\theta^2}{3n} \underset{n\to\infty}{\to} 0$, по достаточному условия оценка состоятельная

Рассматриваем $\tilde{\theta}_2$:

$$M[\tilde{\theta}_2] = \int_{-\infty}^{\infty} y\varphi(y)dy$$

$$\Phi(y) = 1 - (1-F(y))^n \qquad \varphi(y) = n(1-F(y))^{n-1}p(y)$$

$$M[\tilde{\theta}_2] = \int_0^\theta n(1-\frac{y}{\theta})^{n-1}\frac{1}{\theta}ydy =$$

$$t = 1 - \frac{y}{\theta}$$

$$= -\int_1^0 nt^{n-1}(1-t)\theta dt = \int_0^1 n\theta t^{n-1}dt - \int_0^1 n\theta t^n dt =$$

$$= n\theta[1 - \frac{n}{n+1}] = \frac{\theta}{n+1} \quad \text{смещённая}$$

$$\tilde{\theta}_2' = (n+1)x_{min} = (n+1)\tilde{\theta}_2 \quad \text{несмещённая}$$

$$M[\tilde{\theta}_2^2] = \int_0^\theta n(1-\frac{y}{\theta})^{n-1}\frac{1}{\theta}y^2dy =$$

$$= -\int_1^0 nt^{n-1}(1-t)^2\theta^2 dt = \frac{2\theta^2}{(n+1)(n+2)}$$

$$D[\tilde{\theta}_2] = \frac{2\theta^2}{(n+1)(n+2)} - \frac{\theta^2}{(n+1)^2} = \theta^2\left[\frac{2(n+1)-(n+2)}{(n+1)^2(n+2)}\right] =$$

$$= \theta^2\left[\frac{n}{(n+1)^2(n+2)}\right] \underset{n\to\infty}{\to} 0 \quad \text{достаточное не выполнятеся}$$

$$D[\tilde{\theta}_2'] = (n+1)^2 D[\tilde{\theta}_2] = \frac{\theta^2 n}{n+2} \not\to 0$$

Смотрим состоятельность $\tilde{\theta}_2'$ по определению

$$\forall\theta > 0\, \forall\varepsilon > 0 \hookrightarrow P(|\tilde{\theta}_2' - \theta| \geq \varepsilon) \underset{n\to\infty}{\to} 0$$

$$P(|\tilde{\theta}_2' - \theta| \geq \varepsilon) \geq P(\tilde{\theta}_2' > \theta + \varepsilon) =$$

$$= P((n+1)x_{min} \geq \theta + \varepsilon) = P(x_{min} \geq \frac{\theta+\varepsilon}{n+1}) =$$

$$= 1 - P(x_{min} < \frac{\theta+\varepsilon}{n+1}) = 1 - (1 - (1 - F(\frac{\theta+\varepsilon}{n+1}))^n) =$$

$$= (1 - (\frac{\theta+\varepsilon}{\theta(n+1)}))^n \underset{n\to\infty}{\to} e^{-\frac{\theta+\varepsilon}{\theta}} > 0$$

Не является состоятельной!

Смотрим состоятельность $\tilde{\theta}_2$ по определению:

$$P(\tilde{\theta}_2 < \theta - \varepsilon) + \underbrace{P(\tilde{\theta}_2 > \theta + \varepsilon)}_{=0,\text{т.к.}\tilde{\theta}_2 = x_{min}}$$

$$P(x_{min} < \theta - \varepsilon = \Phi(\theta - \varepsilon)) = 1 - (1 - \frac{\theta - \varepsilon}{\theta})^n = 1 - \left(\frac{\varepsilon}{\theta}\right)^n \underset{n\to\infty}{\to} 1$$

Не является состоятельной!

Рассматриваем $\tilde{\theta}_3 = x_{max}$:

$$M[\tilde{\theta}_3] = \int_{-\infty}^{+\infty} z\psi(z)dz = \int_0^\theta n\frac{z^n}{\theta^n}dz = \frac{n}{n+1}\theta \qquad \text{смешённая}$$

$$\Psi(z) = (F(z))^n \qquad \psi(z) = n(F(z))^{n-1}p(z) = n\left(\frac{z}{\theta}\right)^{n-1}\frac{1}{\theta}\{(0;\theta)\}$$

$$D[\tilde{\theta}_3]\frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \frac{n\theta^2}{(n+2)(n+1)^2}$$

$$D[\tilde{\theta}_3']\frac{(n+1)^2}{n^2}D[\tilde{\theta}_3] = \frac{\theta^2}{n(n+2)} \underset{n\to\infty}{\to} 0 \qquad \text{состоятельная}$$

Смотрим состоятельность $\tilde{\theta}_2'$ по определению

$$\forall\theta > 0 \; \forall\varepsilon > 0$$

$$P(|\tilde{\theta}_2' - \theta| \geq \varepsilon) = P(x_{max} < \theta - \varepsilon) + \underbrace{P(x_{max} \geq \theta + \varepsilon)}_{=0} =$$

$$= (F(\theta - \varepsilon))^n = \begin{cases} 0 < \varepsilon < \theta: \; \left(\frac{\theta - \varepsilon}{\theta}\right)^n \underset{n\to\infty}{to} 0 \\ \varepsilon \geq \theta: \; (0)^n \underset{n\to\infty}{\to} 0 \end{cases}$$

Является состоятельной!

Рассматриваем $\tilde{\Theta}_4$:

$$M[\tilde{\theta}_4] = M[x_1 + \sum_{i=2}^n x_i] = M[x_1] + \frac{1}{n-1}\sum_{i=1}^n M[x_i] = \frac{\theta}{2} + \frac{\theta}{2} = \theta$$

$$D[\tilde{\theta}_4] = D[\xi] + \frac{1}{(n-1)^2}(n-1)D[\xi] = \frac{\theta^2}{12}\frac{n}{n-1} \underset{n\to\infty}{\not\to} 0$$

Достаточое усл. не работает.

Используем теорему $\xi_n \xrightarrow{p} \xi$, $\eta_n \xrightarrow{p} \eta$, $\xi_n + \eta_n \xrightarrow{p} \xi + \eta$.

И ЗБЧ Хинчина: $\xi_1, \dots, \xi_n$ незав., одинак распр. $\Rightarrow \frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{p} M[\xi]$.

$$x_1 \xrightarrow{p} x_1 \qquad \frac{1}{n-1}\sum_{i=2}^{n} x_i \xrightarrow{p} \frac{\theta}{2}$$
$$\tilde{\theta}_4 \xrightarrow{p} x_1 + \frac{\theta}{2}$$

Не состоятельна!

Адекватные остались $\tilde{\theta}_1 = 2\bar{x}$, $\tilde{\theta}_3' = \frac{n+1}{n} x_{max}$

$$D[\tilde{\theta}_1] = \frac{\theta^2}{3n} > D[\tilde{\theta}_3'] = \frac{\theta^2}{n(n+2)}$$

Лучшая оценка $\tilde{\theta}_3$.

4.2 Оптимальность и эффективность оценок

Определение 4.1. Несмещённая оценка $\tilde{\theta}(\vec{x}_n)$ характеристики θ называется оптимальной $\tilde{\theta}_{opt}$ если для $\forall \theta \in \Theta \Rightarrow D[\tilde{\theta}_{opt} = \inf D[\tilde{\theta}]$, inf по всем несмещённым оценкам θ.

Теорема 4.2 (Единственность оптимальной оценки)**.** Если оптимальная оценка существует, то она единственна.

Доказательство. Пусть $\tilde{\theta}_1$ и $\tilde{\theta}_2$ разные оптимальные оценки

$$\tilde{\theta}_3 = \frac{\tilde{\theta}_1 + \tilde{\theta}_2}{2} \qquad M[\tilde{\theta}_3] = \theta$$
$$D[\tilde{\theta}_3] = \frac{1}{4} D[\tilde{\theta}_1] + D[\tilde{\theta}_2] + \frac{1}{2} cov(\tilde{\theta}_1, \tilde{\theta}_2) = \frac{1}{2} D[\tilde{\theta}_1] + \frac{1}{2} cos(\tilde{\theta}_1, \tilde{\theta}_2)$$
$$D[a\xi + b\eta] = a^2 D[\xi] + b^2 D[\eta] + 2abcov(\xi, \eta)$$
$$|cov(\tilde{\theta}_1, \tilde{\theta}_2)| \leq \sqrt{D\tilde{\theta}_1 D\tilde{\theta}_2} = D[\tilde{\theta}_1]$$
$$D[\tilde{\theta}_3] \leq D[\tilde{\theta}_1] \qquad D[\tilde{\theta}_3] = D[\tilde{\theta}_1]$$
$$cov(\tilde{\theta}_1, \tilde{\theta}_2) = D[\tilde{\theta}_1] \quad \Rightarrow \quad r = 1 \Leftrightarrow \tilde{\theta}_1 = a\tilde{\theta}_2 + b$$
$$M[\tilde{\theta}_1] = M[\tilde{\theta}_2] = \theta \qquad D[\tilde{\theta}_1] = D[\tilde{\theta}_2]$$
$$\begin{cases} \theta = a\theta + b \\ a^2 D[\tilde{\theta}_2] = D[\tilde{\theta}_2] \end{cases} \Rightarrow \begin{cases} a = 1 \\ b = 0 \end{cases}$$
$$\Rightarrow \tilde{\theta}_1 = \tilde{\theta}_2$$

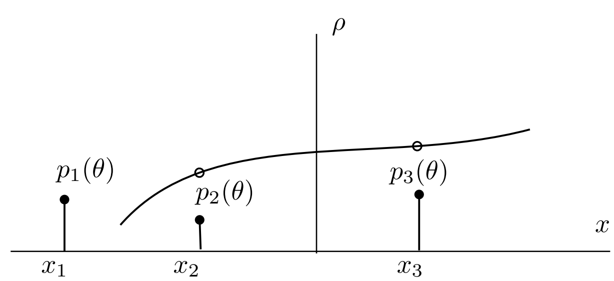
Противоречие. ∎

Будем рассматриватть параметрические вероятностные модели:

$$\xi \sim \rho(x,\theta),\ \theta \in \Theta \subset \mathbb{R},\ x \in A(\theta)$$

$$\xi \sim \rho(x,\vec{\theta}),\ \vec{\theta} \in \Theta \subset \mathbb{R}^m,\ x \in A(\vec{\theta})$$

$$\rho(x,\theta) = \underbrace{p(x,\theta)\{E\}}_{\text{непр. часть}} + \underbrace{\sum_k p_k(\theta)\{x_k\}}_{\text{дискр. часть}}$$



4.3 Информация Фишера

$$I(\theta) = M\left[(\frac{\partial \ln \rho(x,\theta)}{\partial \theta})^2\right] =$$

$$= \int_E \left(\frac{\partial \ln p(x,\theta)}{\partial \theta}\right)^2 p(x,\theta)dx + \sum_k \left(\frac{\partial \ln p_k(x,\theta)}{\partial \theta}\right)^2 p_k(\theta)$$

$I(\vec{\theta})$ - информационная матрица Фишера

$$I_{ij}(\vec{\theta}) = M\left[\frac{\partial \ln p(x,\theta)}{\partial \theta_i}\frac{\partial \ln p(x,\theta)}{\partial \theta_j}\right]$$

Определение 4.2. Вероятностьная модель $\xi \sim \zeta(x,\theta)$, $\theta \in \Theta \subset \mathbb{R}$, $x \in A$ называется регулярной, если

1. $\rho(x,\theta)$ непр дифф по θ на Θ
2. $\frac{\partial}{\partial\theta}\int_A \rho(x,\theta)dx = \int_A \frac{\partial}{\partial\theta}\rho(x,\theta)dx$ на Θ
3. $I(\theta)$ непр на Θ и $I(\theta) > 0$ на Θ

Определение 4.3. Вероятностьная модель $\xi \sim \zeta(x, \vec{\theta})$, $\vec{\theta} \in \Theta \subset \mathbb{R}^m$, $x \in A$ называется регулярной, если

1. $\rho(x, \vec{\theta})$ непр дифф по $\vec{\theta}$ на Θ
2. $\frac{\partial}{\partial \theta_i} \int_A \rho(x, \vec{\theta}) dx = \int_A \frac{\partial}{\partial \theta_i} \rho(x, \vec{\theta}) dx$ на Θ, $i = 1, \ldots, m$
3. $I(\vec{\theta})$ положительно определена на Θ и $I_{ij}(\vec{\theta})$ непрер. на Θ

Определение 4.4. Вероятностная модель $\xi \sim \rho(x, \theta)$, $\vec{\theta} \in \Theta \subset \mathbb{R}^m$, $x \in A$ называется сильно регулярной, если эта модель регулярна и

1. $\rho(x, \theta)$ k раз непрерывно дифф по θ на Θ ($k \geq 2$)
2. $\frac{\partial^l}{\partial \theta^l} \int_A \rho(x, \theta) dx = \int_A \frac{\partial^l}{\partial \theta^l} \rho(x, \theta) dx$, $l = 1, \ldots, k$

Определение 4.5. Вероятностная модель $\xi \sim \zeta(x, \vec{\theta})$, $\vec{\theta} \in \Theta \subset \mathbb{R}^m$, $x \in A$ называется сильно регулярной, если эта модель регулярна и

1. $\rho(x, \vec{\theta})$ k раз непрерывно дифф по θ на Θ ($k \geq 2$)
2. все производные по $\vec{\theta}$ перестановочные с $\int$ по x

Определение 4.6. Статистика $\tilde{g}(\vec{x}_n)$ называется регулярной оценкой функции $g(\theta)$, если она является несмещённой оценкой и

$$\frac{\partial}{\partial \theta} \int_B \tilde{g}(\vec{x}_n) L(\vec{x}_n, \theta) d\vec{x}_n = \int_B \tilde{g}(\vec{x}_n) \frac{\partial}{\partial \theta} L(\vec{x}_n, \theta) d\vec{x}_n$$

где $L(\vec{x}_n, \theta)$ - плотность распределения случайного вектора $\vec{x}_n$ ($L(\vec{x}_n, \theta) = \prod_{i=1}^n \rho(x_i, \theta)$), $B = \underbrace{A \times A \times \cdots \times A}_{n}$

Теорема 4.3 (Достаточное условие регулярности оценки)**.** Если модель регулярна, $\tilde{g}(\vec{x}_n)$ является несмещ. оценкой $g(\theta)$ и $D[\tilde{g}(\vec{x}_n]$ ограничена на $\forall$ компакте из Θ по θ, тогда оценка регулярна.