Математическая статистика

Матвеев Сергей М3338

5 семестр

1 Предмет математической статистики. Задачи, решаемые методами математической статистики.

Пусть у нас есть генеральная совокупность, но мы хотим ее как-то изучать, тогда мы можем взять ее часть - выборку.
Мы хотим по выборке сделать содержательные вероятностные выводы о генеральной совокупности.
Примеры задач, которые могут быть решены таким способом:

1. Бросок монеты (оценить вероятность орла, честно|нечестно)
2. Замеры показателя: какие типичные значения для показателя
3. Как учатся мальчики и девочки (одинаково или по разному)
4. Цена на недвижимость, расстояние до метро (оценка зависимости)

2 Модель простейшей выборки. Эмпирическая функция распределения и её свойства. Способы визуализации выборки.

Простейшая модель выборки - $X_1, X_2, \dots, X_n$ - *i.d.d.*, F - функция распределения (теоретическая функция).
$X_1, \dots, X_n \sim F$ (F мы не знаем априори)
$x_1, \dots, x_n$ - реализация выборки
Цель: оценить из реализации $x_1, \dots, x_n$ теор F.
Эмпирическая фукнция распределения:
$\mu_n(t) = \sum_{i=1}^{n} \mathbb{1}(X_i \le t)$
$F_n(t) = \frac{\mu_n(t)}{n}$ - эмпирическая функция распределения.

Свойства

$\mathbb{1}(X_i \le t) \sim Bern(F(t))$

$\mathbb{E}(F_n(t)) = F(t)$ (это называется несмещенность)

$\mathrm{Var}(F_n(t)) = \dfrac{F(t)(1-F(t))}{n}$

ЗБЧ: $F_n(t) \xrightarrow{P} F(t)$ - это называется состоятельность

ЦПТ: $\dfrac{\mu_n(t) - nF(t)}{\sqrt{F(t)(1-F(t))n}} \xrightarrow{d} U \sim N(0,1) =$

$= \sqrt{n}\dfrac{F_n(t) - F(t)}{\sqrt{F(t)(1-F(t))}}$ (ассимптотическая нормальность)

Теорема Гливенко-Кантелли

$\sup\limits_{t\in\mathbb{R}}|F_n(t) - F(t)| \xrightarrow[n\to\infty]{a.s.} 0$

Теорема Колмогорова

$$D_n = \sup_x |F_n(x) - F(x)| \Rightarrow P(\sqrt{n}D_n \le t) \xrightarrow[n\to\infty]{} K(t) = \sum_{j=-\infty}^{+\infty} (-1)^j e^{-2j^2t^2}$$

$F \in C(\mathbb{R})$

Такая функция называется функцией Колмогорова

Теорема Смирнова

$X_1, \dots, X_n$, $Y_1, \dots, Y_n$ - независимы

Обе распределены по $F \in C(\mathbb{R})$

$$D_{n,m} = \sup_x |F_n(x) - F_m(x)| \Rightarrow P(\sqrt{\frac{mn}{m+n}}D_{n,m} \le t) \xrightarrow[n\to\infty, m\to\infty]{} K(t)$$

Стоит отметить, что обе теоремы имеют смысл при $t \ge 0$

Выборку можно визуализировать используя гистограмму или box plot

3 Начальные выборочные моменты и их свойства, в том числе выборочное среднее.

$\alpha_k = EX_1^k$ - к-ый теоретический момент.

$\overline{g(X)} = \dfrac{1}{n}\sum\limits_{k=1}^{n} g(X_k)$, $g : \mathbb{R} \to \mathbb{R}$

$\widehat{\alpha_k} = \overline{X^k} = \dfrac{1}{n}\sum\limits_{j=1}^{n} X_j^k$ - к-ый выборочный момент.

$E\widehat{\alpha_k} = \alpha_k$ (несмещенность, мы просто воспользовались линейностью математического ожидания)

$\mathrm{Var}\,\widehat{\alpha_k} = \dfrac{1}{n}\mathrm{Var}(X_1^k) = \dfrac{1}{n}(EX_1^{2k} - (EX_1^k)^2)$

По ЦПТ получаем:

$\sqrt{n}\dfrac{\widehat{\alpha_k} - \alpha_k}{\sqrt{\alpha_{2k} - \alpha_k^2}} \approx N(0,1)$

$\sqrt{n}\frac{\widehat{\alpha_k}-\alpha_k}{\sqrt{\widehat{\alpha_{2k}}-\widehat{\alpha_k}^2}} = \sqrt{n}\frac{\widehat{\alpha_k}-\alpha_k}{\sqrt{\alpha_{2k}-\alpha_k^2}} \cdot \sqrt{\frac{\alpha_{2k}-\alpha_k^2}{\widehat{\alpha_{2k}}-\widehat{\alpha_k}}}$ - первый множитель по ЦПТ сходится к $N(0,1)$ по распределнию

Давайте посмотрим что будет со вторым множителем. Он будет сходиться к 1 по вероятности.

Таким обрбазом:

$\sqrt{n}\frac{\widehat{\alpha_k}-\alpha_k}{\sqrt{\alpha_{2k}-\alpha_k^2}} \cdot \sqrt{\frac{\alpha_{2k}-\alpha_k^2}{\widehat{\alpha_{2k}}-\widehat{\alpha_k}}} \xrightarrow{d} N(0,1)$

А почему вторая дробь сходится к единице?

$\widehat{\alpha_k} \xrightarrow{P} \alpha_k$ (по ЗБЦ, это называется состоятельность)

$\widehat{\alpha_{2k}} - \widehat{\alpha_k}^2 \xrightarrow{P} \alpha_{2k} - \alpha_k^2$

$\overline{X}$ - выборочное среднее.

4 Выборочные центральные моменты и их свойства, в том числе выборочная дисперсия. Дельта-метод.

$\beta_k = E(X_1 - EX_1)^k$ - к-ый центральный момент.

$\widehat{\beta_k} = \overline{(X-\overline{X})^k} = \frac{1}{n}\sum_{j=1}^{n}(X_j-\overline{X})^k$ - к-ый выборочный момент.

$\widehat{\beta_2} = S_*^2$ - выборочная дисперсия.

S_* - выборочное стандартное отклонение (выборочное среднеквадратичное отклонение).

$Note$: выборочные моменты есть ничто иное как моменты посчитанные относительно эмпирического распределения.

$S_* = \overline{X^2} - (\overline{X}^2)$

$\widehat{\beta_k} = Poly(\widehat{\alpha_k}, \dots, \widehat{\alpha_1})$

$\widehat{\alpha_1}, \dots, \widehat{\alpha_k}$ - состоятельные оценки (имеет место сходимость по вероятности)

Так как полином это непрерывная функция, то $\widehat{\beta_k} \xrightarrow{P} \beta_k$

$ES_*^2 = \frac{1}{n}\sum_{i=1}^{n} EX_i^2 - E(\overline{X})^2 = \alpha_2 - \alpha_1^2 - \frac{\beta_2}{n} = \beta_2 - \frac{\beta_2}{n} = \frac{n-1}{n}\beta_2$

Т.к. $ES_*^2 \neq \beta_2$ такую дисперсию называют несмещенной

$S^2 = \frac{n}{n-1}S_*^2$ - несмещенная дисперсия

5 Выборочные квантили, в том числе выборочная медиана. Распределение k-ой порядковой статистики.

Определение. Вариационный ряд
$X_{(1)} \leq X_{(3)} \leq \cdots \leq X_{(n)}$ - вариационный ряд
Определение. Порядковая статистика
$X_{(k)}$ - к-я порядковая статистика.
Квантиль порядка α
$P(X \geq q_\alpha) \geq 1 - \alpha$
$P(X \leq \alpha) \geq \alpha$
Это общее определение
Если F строго возрастает:
$F(q_\alpha) = \alpha \Leftrightarrow q_\alpha = F^{-1}(\alpha)$
$F^{-1}(\alpha) : \sup\{x : F(x) \leq \alpha\}, \inf\{x : F(x) \geq \alpha\}$
Определение. Выборочный квантиль порядка α
$\alpha \in (0, 1)$
$\exists 0 \leq k \leq n-1 : \dfrac{k}{n} \leq \alpha < \dfrac{k+1}{n}$
$X_{(k+1)}$ - выборочный квантиль порядка α
$\alpha = 0$ $\min(X)$ - нулевой квартиль
$F^{-1} = \sum\{x \in \mathbb{R} F_n(x) \leq \alpha\}$
$\alpha = \dfrac{1}{4}$ - первый вартиль (нижний квартиль)
$\alpha = \dfrac{1}{2}$ - второй квартиль (выборочная медиана)
$\alpha = \dfrac{3}{4}$ - третий квартиль (верхний квартиль)
$\alpha = 1$ - $\max(X)$ (четвертый квартиль)
$n = 2m \Rightarrow = med(X) = \dfrac{X_{(m)} + X_{(m+1)}}{2}$
$n = 2m + 1 \Rightarrow med(X) = X_{(m+1)}$
$IQR = \Delta$ между верхним и нижним квартилем.
$$P(X_{(k)} \leq t) = P(\mu_n(t) \geq k) = \sum_{j=k}^{n} C_n^j F^j(t)(1 - F(t))^{n-j}$$
$$B(z, a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^z t^{a-1}(1-t)^{b-1} dt = B(F(t), k, n-k+1)$$
$0 \leq z \leq 1$
Пусть $p()$ - теоретическая плотность, то есть $p = F'$
$(P(X_{(k)} \leq t))'_t = \dfrac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} \cdot F^{k-1}(t)(1 - F(t))^{n-k} \cdot p(t)$ - плотность к-й порядковой статистики

6 Теоремы об асимптотиках среднего и крайнего членов вариационного ряда (идеи доказательств).

Средний член вариационного ряда: $\frac{K(n)}{n} \to const \in (0,1)$
Крайний член вариационного ряда:
$X_{(r)}$, r - огр.
$X_{(n+1-s)}$, s - огр.
Теорема об ассимптотике среднего члена вариационного ряда
$0 < \alpha < 1$ - теоретическая плотность.
q_α - теоретический квантиль порядка α
$p \in C^1$(окр-сть q_α)
$p(q_\alpha) > 0$
Тогда:
$\sqrt{n} \cdot f(q_\alpha)\frac{X_{(\lfloor n\alpha \rfloor)} - q_\alpha}{\sqrt{\alpha(1-\alpha)}} \xrightarrow{d} N(0,1)$
Идея доказательства
Пусть $\lfloor n\alpha \rfloor = k$
Мы умеем писать плотность для $X_{(k)}$
Затем у нас идет преобразование:
$g(x) = \sqrt{n}p(q_\alpha)\frac{x - q_\alpha}{\sqrt{\alpha(1-\alpha)}} \rightsquigarrow p_{g(X_{(k)})}(t) = p_{X_{(k)}}(g^{-1}(t))|g^{-1}(t)'_t|$ (теорема из прошлого семестра)
Там вылезут факториалы, от них мы умеем избавляться по Стирлингу
Затем надо будет воспользоваться непрерывной дифференцируемостью:
$p \in C^1$(окр-сть q_α)
$p(q_\alpha) > 0$
Тогда в пределе наша новая плотность будет стремиться к плотности нормального стандартного закона.
Теорема об ассимптотике крайних членов вариационного ряда
$r, s, F, x, p()$ - плотность
Тогда:
$nF(X_{(r)}) \xrightarrow{d} \Gamma(r,1)$
$n(1 - F(X_{(n+1-s)})) \xrightarrow{d} \Gamma(s,1)$
И оба распределения независимы.
Идея доказательства
У нас есть совместная плотность и какое-то преобразование, тогда мы можем написать плотность после преобразования
Затем берем предел и мы получим плотность равная произведению двух этих двух законов.

7 Постановка задачи точечного оценивания параметров. Свойства оценок.

$X_1, \dots, X_n \sim F_\theta \in \Theta \subset \mathbb{R}^d$
θ - некий фиксированный неизвестный вектор.
Наша цель оценить θ в виде $\widehat{\theta} = \widehat{\theta}(X_1, \dots, X_n)$
Определение. Состоятельность
$\widehat{\theta}$ - состоятельная оценка $\theta \Leftrightarrow \widehat{\theta} \xrightarrow{P} \theta$
Определение. Несмещенность
$b.as(\widehat{\theta}) \stackrel{def}{=} E\widehat{\theta} - \theta$ - смещение
$b.as(\widehat{\theta}) = 0 \Leftrightarrow$ несмещенная
Определение. Ассимптотическая нормальность
$\sqrt{n}(\widehat{\theta} - \theta) \to N(0, \Sigma_\theta)$
Определение. Эффективность (оптимальность)
$\widehat{\theta_1}$ эффективнее $\widehat{\theta_2} \Leftrightarrow MSE\widehat{\theta_1} < MSE\widehat{\theta_2}$
$MSE\widehat{\theta} = E\left\|\widehat{\theta} - \theta\right\|^2 = E(\widehat{\theta} - \theta)^T(\widehat{\theta} - \theta)$
Утверждение
$MSE\widehat{\theta} = tr(\operatorname{Var}\widehat{\theta}) + \left\|b.as\widehat{\theta}\right\|^2$
Доказательство
$MSE = E(\widehat{\theta} - \theta)^T(\widehat{\theta} - \theta) = E(\widehat{\theta} - E\widehat{\theta} + E\widehat{\theta} - \theta)^T(\widehat{\theta} - E\widehat{\theta} + E\widehat{\theta} - \theta) = E(\widehat{\theta} - E\widehat{\theta})^T(\widehat{\theta} - E\widehat{\theta}) = \sum \operatorname{Var}\widehat{\theta_i} + \left\|b.as\widehat{\theta}\right\|^2$

1. Ассимптотическая нормальность $\Rightarrow$ состоятельность
   $\widehat{\theta} - \theta = \frac{1}{\sqrt{n}}\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{P} 0$

2. Ассимптотическая нормальность $\Rightarrow b.as\widehat{\theta} \to 0$
   Пусть $d = 1$
   $P(|\theta - E\widehat{\theta}| > \varepsilon) = P(\frac{\sqrt{n}|\theta - E\widehat{\theta}|}{\sigma} > \frac{\varepsilon\sqrt{n}}{\sigma}) = 1 - P(\cdots < \frac{\varepsilon\sqrt{n}}{\sigma}) \approx 1 - (2\Phi(\frac{\varepsilon\sqrt{n}}{\sigma}) - 1) = 2(1 - \Phi(\frac{\varepsilon\sqrt{n}}{\sigma})) \to 0$

3. Состоятельность $\Rightarrow b.as\widehat{\theta} \to 0$
   Следует из усиленного закона больших чисел
   $\overline{X} \xrightarrow{a.s} \mu \Rightarrow E\overline{X} \to \mu$ (По теореме Лебега о мажорируемой сходимости)

4. Пусть $d = 1$, $b.as\widehat{\theta} \to 0, \operatorname{Var}\widehat{\theta} \to 0 \Rightarrow \widehat{\theta}$ - сост.

8 Метод моментов.

Рассмотрим $g_1, \dots, g_d$
$\exists Eg_1(X_1) = m_1(\theta_1, \dots, \theta_d)$

$\exists Eg_2(X_2) = m_2(\theta_1, \dots, \theta_d)$
$\dots$
$\exists Eg_d(X_d) = m_d(\theta_1, \dots, \theta_d)$
$\begin{cases} \overline{g_1(X)} = m_1(\widehat{\theta_1}, \dots, \widehat{\theta_d}) \\ \dots \\ \overline{g_d(X)} = m_d(\widehat{\theta_1}, \dots, \widehat{\theta_d}) \end{cases}$
Пусть $\exists!$ решение:
$\begin{cases} \widehat{\theta_1} = \alpha_1(\overline{g_1(X)}, \dots, \overline{g_d(X)}) \\ \dots \\ \widehat{\theta_d} = \alpha_d(\overline{g_1(X)}, \dots, \overline{g_d(X)}) \end{cases}$
Тогда это будет оценка методов моментов.

9 Метод максимального правдоподобия.

probability must function: $p(x, \theta) = p(x|\theta)$
probability identity function: $p(x, \theta) = p(x|\theta)$
Будем называть оба случая плотностью.
Пусть у нас есть выборка $X_1, \dots, X_n \sim p(x|\theta)$
$L(x|\theta) = \prod p(x_i|\theta)$ - функция правдоподобия
$\theta^* = \underset{\widehat{\theta}}{argmax}(L(x, \theta))$ - оценка максимума правдоподобия
$\theta \in \Theta$ - открыто
$\theta_1 \neq \theta_2 \Rightarrow L(x, \theta_1) \neq L(x, \theta_2)$
Доказательство

1. Посмотреть и подумать
2. Рассмотреть $\ln L(x|\theta)$; $\dfrac{\partial \ln L(x, \theta)}{\partial \theta}$
3. $\dfrac{\partial \ln L(x, \theta)}{\partial \theta} = 0$
4. Проверить достаточные условия максимума

10 Оценки максимального правдоподобия для нормальной и полиномиальной моделей.

1) $N(\theta_1, \theta_2)$
$L(x, \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} \exp\left(-\frac{(x_i - \theta_1)^2}{2\theta_2}\right)$
$\ln L(x, \theta) = \sum_{i=1}^{n} [-\frac{1}{2} \ln 2\pi - \frac{1}{2}\theta_2 - \frac{(x_i - \theta_1)^2}{2\theta_2}]$

$$\frac{\partial \ln L(x,\theta_1)}{\theta_1} = \sum_{i=1}^{n} \frac{2(x_i-\theta_1)}{2\theta_2} = \sum_{i=1}^{n} \frac{x_i-\theta_1}{\theta_2}$$
$$\frac{\partial \ln L(x,\theta)}{\partial \theta_2} = \sum_{i=1}^{n} [-\frac{1}{2\theta_2} + \frac{(x_i-\theta_1)^2}{2\theta_2^2}]$$
$$\sum_{i=1}^{n} \frac{(x_i-\widehat{\theta_1})}{\widehat{\theta_2}} = 0 \Rightarrow \widehat{\theta_1} = \overline{X}$$
$$\sum_{i=1}^{n} [-\frac{1}{2\widehat{\theta_2}} + \frac{(x_i-\widehat{\theta_1})^2}{2\widehat{\theta_2}^2}] = 0 \Rightarrow \widehat{\theta_2} = S_*^2$$
2) $Poly(1,p) : p = (p_1,\dots,p_m)$
Рассмотрим частоты:
ν_1 - кол-во наблюдений типа 1
$\dots$
ν_m - кол-во наблюдений типа m
Суммируем и смотрим на функцию правдоподобия
$L(X,p) = p_1 \dots p_m$
$$\ln L(X,p) = \sum_{j=1}^{m-1} \nu_j \ln p_j + \nu_m \ln(1-p_1-\dots-p_{m-1})$$
$$\frac{\partial \ln L \dots}{\partial p_j} = \frac{\nu_j}{p_j} - \frac{\nu_m}{1-p_1-\dots-p_{m-1}} = 0$$
$\sum$ уравнения: $\nu_j(1-\widehat{p_1}-\dots-\widehat{p_{m-1}}) = \widehat{p_j}\cdot\nu_m$
$\widehat{p_m}(n-\nu_m) = \nu_m(1-\widehat{p_m})$
$\widehat{p_m}n - \widehat{p_m}\nu_m = \nu_m$
$\widehat{p_m} = \frac{\nu_m}{n}$
$\widehat{p_j} = \frac{\nu_j\widehat{p_m}}{\nu_m} = \frac{\nu_j}{n}$

11 Информация Фишера.

$d=1 : L(X,\theta) = \prod p(X_j,\theta)$
$\ln L(X,\theta) = \sum \ln p(x_j,\theta)$
$V(X,\theta) = \frac{\partial \ln L \dots}{\partial\theta} = \sum \frac{\partial \ln p \dots}{\partial\theta}$ - вклад выборки
$\theta \in \Theta$ - открыто
$\theta_1 \neq \theta_2 \Rightarrow p(X,\theta_1) \neq p(X,\theta_2)$
Регулярность:

1. $\frac{\partial}{\partial\theta}\int_X T(X)L(X_i,\theta)dX = \int \frac{\partial}{\partial\theta}L(X,\theta)\cdot T(X)dX$
   Необходимое условие $\sup p_x$ не зависит от θ
   $U[0,\theta]$ $\int_0^\theta \frac{1}{\theta}dt = 1$

$$(\int_0^\theta \frac{1}{\theta}dt)'_\theta = (\frac{1}{\theta}\int_0^\theta dt)'_\theta = -\frac{1}{\theta^2}\int_0^\theta dt + \frac{1}{\theta} = 0 \neq \int_0^\theta (\frac{1}{\theta})'_\theta dt$$

2. $EV^2(X,\theta) < \infty$

$\int_X L(X,\theta)dX = 1 \xrightarrow{\frac{\partial}{\partial\theta}} \int_X \frac{\partial L(.)}{\partial\theta}dX = \int_X \frac{\frac{\partial L(\dots)}{\partial\theta}}{L(\dots)} \cdot L(\dots)dX = \int_X V(X,\theta)L(X,\theta)dX =$
$EV(X,\theta) = 0$
$I(\theta) = \mathrm{Var}(V(X_i,\theta)) = E(V^2(X_i,\theta))$ - информация Фишера для всей выборки
$V(X,\theta) = \sum_j \frac{\partial \ln p(x,\theta)}{\partial\theta} \Rightarrow \mathrm{Var}(V(X,\theta)) = n \cdot \mathrm{Var}\,\frac{\partial \ln p(x,\theta)}{\partial\theta}$
$i(\theta)$ - информация Фишера для 1 наблюдения
$i(\theta) = E(\frac{\partial \ln p(x_j,\theta)}{\partial\theta})^2$
$\frac{\partial}{\partial\theta}\int_{\mathbb{R}} \frac{\partial \ln p(x,\theta)}{\partial\theta} \cdot p(x,\theta)dx = \int_{\mathbb{R}} \frac{\partial^2 \ln p(x,\theta)}{\partial^2\theta}dx + \int_{\mathbb{R}} \frac{\partial \ln \dots}{\partial\theta}\frac{\partial p \dots}{\partial\theta}$вот тут домножаем и делим на плотность
$E\frac{\partial^2 \ln p(x,\theta)}{\partial\theta^2} + E(\frac{\partial \ln p(x,\theta)}{\partial\theta})^2 = 0$
$i(\theta) = E(\frac{\partial \ln p(x_j,\theta)}{\partial\theta})^2 = -E\frac{\partial^2 \ln p(x,\theta)}{\partial\theta^2}$
Произвольное d:
$i(\theta) = -(E\frac{\partial^2 \ln p(X,\theta)}{\partial\theta_i\partial\theta_j})_{1\le i,j\le d}$
$I(\theta) = ni(\theta)$

12 Неравенство Рао-Крамера.

Модель регулярная, d = 1
$\tau(\theta)$ - оцениваемая функция
$\tau \in C^1$ (как правило $\tau(\theta) = \theta$)
$E\widehat{\tau(\theta)} = \theta$
Тогда:
$\mathrm{Var}\,\widehat{\tau(\theta)} \ge \frac{[\tau'(\theta)]^2}{ni(\theta)}$
$\tau'(\theta) = \int \widehat{\tau(\theta)}\frac{\partial L(X,\theta)}{\partial\theta}dX = \int \widehat{\tau(\theta)}V(X,\theta)L(X,\theta)dX - EV(X,\theta)\cdot E\widehat{\tau(\theta)} =$
$\mathrm{Cov}(V(X,\theta),\widehat{\tau(\theta)})$
$\mathrm{Cov}^2(V(X,\theta),\widehat{\tau(\theta)}) \le \mathrm{Var}(V(X,\theta)) \cdot \mathrm{Var}(\widehat{\tau(\theta)})$
Многомерный случай
$\tau(\theta) : \mathbb{R}^d \to \mathbb{R}$
$\tau \in C^1$
$E\widehat{\tau\theta} = \tau\theta \Rightarrow \mathrm{Var}\,\widehat{\tau(\theta)} \ge \frac{\nabla\tau(\theta)i^{-1}(\theta)\nabla^T\tau(\theta)}{n}$

13 Свойства оценок максимального правдоподобия.

Состоятельность

Пусть θ_0 - реальный параметр $\Rightarrow p_{\theta_0}(L(X,\theta_0) > L(X,\theta)) \to 1$

$\frac{L(X,\theta)}{L(X,\theta_0)} < 1$

$\frac{1}{n}\sum \ln \frac{p(X_j,\theta)}{p(X_j,\theta)} < 0$

По ЗБЧ $\Rightarrow E_{\theta_0} \ln \frac{p(x_j,\theta)}{p(x_j,\theta_0)} \le E_{\theta_0}[\frac{p(X_j,\theta)}{p(X_j,\theta_0)} - 1] = \int_X p(X,\theta)dX - \int p(X,\theta_0)dX = 0$

Давайте введем события:

$S_n = \{X : \ln L(X,\theta_0) > \ln L(X,\theta_0 - a)\} \cap \{X : \ln L(X,\theta_0) > \ln L(X_i\theta_0 + a)\}$

$P_{\theta_0}(S_n) \to 1$

$A_n = \{X : |\widehat{\theta} - \theta_0| < a\}$

$B_n = \{X : \frac{\partial \ln L(X,\theta)}{\partial \theta}|_{\theta=\widehat{\theta}} = 0\}$

$S_n \subset A_n B_n \subset A_n \Rightarrow P(A_n) \to 1$

Принцип инвариантности

$\theta \in \Theta \xrightarrow[\varphi]{biection} \gamma \in \Gamma$

$\theta = \varphi^{-1}(\theta) \Leftrightarrow \gamma = \varphi(\theta)$

$\sup\limits_{\theta} L(X,\varphi(\gamma)) = \sup\limits_{\gamma} L(x,\gamma)$

$\gamma* = \varphi(\theta*)$

Теорема Ассимптотическая нормальность ОМП

Пусть наша модель регулярная, так же пусть:

$|\frac{\partial^3 \ln f(x,\theta)}{\partial\theta_i \partial\theta_j \partial\theta_k}| \le M$

θ_* - ОПМ для θ

Уравнение $\nabla \ln L(X,\theta) = 0$ имеет еддинственное решение. Тогда:

1. $\sqrt{n}(\theta_* - \theta) \to N(0, i^{-1}(\theta))$
2. $\tau(\theta)$ - оцениваемая функция от θ
   $\tau \in C^1$
   $\sqrt{n}(\tau(\theta_*) - \tau(\theta)) \to N(0,\sigma^2)$
   $\sigma^2 = \nabla\tau(\theta) i^{-1}(\theta) \nabla^T \tau(\theta)$
3. σ^2 - непрерывная функция от $\theta \Rightarrow \sqrt{\frac{\tau(\theta_*) - \tau(\theta)}{\sigma(\theta_*)}} \to N(0,1)$

14 Экспоненциальное семейство распределений.

Пусть наше распределение относится к экспоненциальному семейству распределений если:
$p(x,\theta) = \exp(A(\theta)B(x) + C(\theta) + D(x))$
К таким распределениям относятся: $N(), \Gamma(), Pois(), Bin, NB$
$\ln p(x,\theta) = A(\theta)B(x) + C(\theta) + D(x)$
$\frac{\partial \ln p(x,\theta)}{\partial \theta} = A'(\theta)B(x) + C'(\theta)$
$V(X,\theta) = A'(\theta)\sum B(X_i) + nC'(\theta)$
$V(X,\theta) = n(A'(\theta)\overline{B(X)} + C''(\theta))$
$\frac{V(X,\theta)}{n} - C'(\theta) = A'(\theta)\overline{B(X)}$
$\overline{B(X)} = \frac{V(X,\theta)}{nA'(\theta)} - \frac{C'(\theta)}{A'(\theta)}$
$\overline{B(X)}$ - оптимальная оценка для $(-\frac{C'(\theta)}{A'(\theta)})$

15 Байесовские оценки.

$X_1, \ldots, X_n \sim F_\theta$
$\theta \sim \pi(\theta)$ - prior
$l(\widehat{\theta}, \theta)$ - функция потерь
$l(\widehat{\theta}, \theta) = (\widehat{\theta} - \theta)^2$ (default)
$R(\widehat{\theta}, \theta) = El(\widehat{\theta}, \theta)$ - риск
$r(\widehat{\theta}) = E_{\pi(\theta)}R(\widehat{\theta}, \theta)$ - байесовский риск
$\widehat{\theta}_B = \underset{\theta}{argmin} r(\widehat{\theta})$
$r(\widehat{\theta}) = El(\widehat{\theta}, \theta)$
Давайте вспомним теорему Байеса:
$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
$P(\theta|X) = \frac{L(X|\theta)\pi(\theta)}{\int L(X|\theta)\pi(\theta)d\theta}$
posterior = likelihoox × prior
$\widehat{\theta}_B = \underset{\widehat{\theta}}{argmin} E[l(\widehat{\theta})|X]$
$r(\theta_*) \leq r(\widehat{\theta})$

16 Минимаксные оценки.

$m(\widehat{\theta}) = \sup\limits_{\theta} R(\widehat{\theta}, \theta)$
$\widehat{\theta_{WC}} = argminm(\widehat{\theta})$ - минимаксная оценка
$r(\widehat{\theta}) \leq m(\widehat{\theta})$
Утверждение
$\exists \pi(\theta)$ - prior : $R(\widehat{\theta_B}, \theta) = \text{const} \Rightarrow \widehat{\theta}_{WC} = \widehat{\theta}_B$

17 Доверительные интервалы. Общая схема построения доверительного интервала. «Универсальный» рецепт.

Определение. Доверительный интервал
$X_1, \ldots, X_n \sim F_\theta, \theta \in \Theta \subset \mathbb{R}$
$1 - \alpha = \gamma \in (0, 1)$ - уровень доверия
default: 0.9, 0.95, 0.99
$(T_l(X), T_r(X))$ - доверительный интервал уровня $\gamma = 1 - \alpha$ если $p(\theta \in (T_l(X), T_r(X)) \geq \gamma)$
Пусть $T(X, \theta) \sim G$ - не зависит от θ
Рассмторим $p(q_1 < T(X, \theta) < q_2) = 1 - \alpha$
$q_1 = q_{\frac{\alpha}{2}}$
$q_2 = q_{1-\frac{\alpha}{2}}$
Из данной вероятности можно выразить θ
Универсальный рецепт (нет)
а) $F_\theta(X_k)$
$P(F_\theta(X_k) \leq t) = P(X_k \leq F_\theta^{-1}(t)) = F_\theta(F_\theta^{-1}(t)) = t$
б) $-\ln F_\theta(X_k) \sim Exp(1)$
$P(-\ln U \leq t) = P(U \geq e^{-t}) = 1 - e^{-t}$
в) $-\sum \ln F_\theta(X_k) \sim \Gamma(n, 1)$

18 Теорема Фишера и примыкающие к ней леммы. Распределение хи-квадрат.

Лемма о независимости линейной и квадратичной статистик
$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$
$T = AX, X = (X_1, \ldots, X_n)^T A \in M_{m \times n}(\mathbb{R})$
$Q = X^T BX, B \in M_n(\mathbb{R}), B = B^T$
$AB = 0$
Тогда T, Q - независимы
Доказательство

$\Lambda = U^T BU$
$\Lambda = diag(\lambda_1, \dots, \lambda_m, 0, 0)$
λ_k - собственное число не 0
$U = (u_1, \dots, u_n)$ - собственные векторы ортонормированного базиса $\Leftrightarrow B = U\Lambda U^T = \sum_{j=1}^{m} \lambda_j u_j uj^T \Rightarrow Q = \sum_{j=1}^{M} \lambda_j (X^T U_j)(U_j^T X) = \sum_j \lambda_j (U_j^T X)^2$
$A(\sum_{j=1}^{m} \lambda_j u_j u_j^T) = 0$
$\sum_{j=1}^{m} \lambda_j A U_j u_j^T = 0$
Зафиксируем $1 \le k \le m$ домножим справа на u_k
$Au_k = 0 \Rightarrow \forall i A[i, *]u_k = 0$
Нам надо доказать, что $\forall i, k A[i, *]X$ и $u_k^T X$ - нез
$\mathrm{Cov}(A[i, *]X, u_k^T X) = \mathrm{Cov}(A[0, *]X, X^T u_k) = A[i, *] \operatorname{Var} X u_k = \sigma^2 A[i, *]u_k = 0$
Лемма о независимости двух квадратичных статистик
$Q_1 = X^T B_1 X$
$Q_2 = X^T B_2 X$
$B_1 B_2 = B_2 B_1 = 0$
Тогда Q_1, Q_2 - нез
Определение X_n - квадратичная
$X_1, \dots, X_n \sim N(0, 1)$
$\sum_{k=1}^{n} X_k^2 \sim \Xi^2(n)$ (распределение)
хи-квадрат с n степенями свободы
Лемма о распределении квадратичной статистики
$X_1, \dots, X_n \sim N(0, 1)$
$Q = X^T BX$
$B = B^2$
Тогда $Q \sim \Xi^2(r), r = rank(B) = tr(B)$
$Q = \sum_{k=1}^{n} (u_k^T X)^2 \sim \Xi^2(r)$
Доказательство
$u_k^T \sim N(u_k^T EX, u_k^T I_n u_k) = N(0, 1)$
$\mathrm{Cov}(u_k^T X, u_j^T X) = 0$
$B = U\Lambda U^T$
$rankB = rank\Lambda = tr\Lambda$
$trB = tr(U\Lambda U^T)$
Заметим что $B_{j,j} = \lambda_j u_j u_j^T = \lambda_j$
Теорема Фишера
$X_1, \dots, X_n \sim N(\mu, \sigma^2) \Rightarrow$

1. $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$
2. $\frac{nS_*^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \Xi^2(n-1)$
3. $S^2, \overline{X}$ - нез
4. $S_*^2, \overline{X}$ - нез

$Y_j = \frac{X_j - \mu}{\sigma}$
$\overline{Y} = \frac{1}{\sigma}(\overline{X} - \mu)$
$S_*^2(Y) = \frac{1}{n}\sum_{j=1}^{n}(Y_j - \overline{Y})^2 = \frac{S_*^2(X)}{\sigma^2}$
$\overline{Y} = \frac{\sum Y_j}{n} = (\frac{1}{n}, \dots, \frac{1}{n})(= b)(Y_1, \dots, Y_n)^T = bY$
$nS_*^2(Y) = (Y - BY)^T(Y - BY) = Y^T(I - B)^T(I - B)Y = \sum \Xi^2(tr(I - B))$ - по предыдущей лемме
Для того чтобы доказать третье утверждение
$b(I - B) = b - b = 0$
Тогда мы пользуемся первой леммой
Таким образом теорема Фишера доказана.

19 Распределения Фишера, Стьюдента. Построение доверительных интервалов для параметров нормального закона.

Определение. Распределение Стьюдента
$X_0, X_1, \dots, X_n$ - нез, $N(0,1)$
$$\frac{X_0}{\sqrt{\frac{1}{n}\sum_{k=1}^{n} X_k^2}} \sim T(n)$$
n - степени свободы (deg of freedom)
Давайте выведем статистику:
$$\frac{\sqrt{n}\frac{\overline{X}-\mu}{\sigma}}{\sqrt{\frac{1}{n-1}\frac{ns_*^2}{\sigma^2}}} = \sqrt{n-1}\frac{\overline{X} - \mu}{S_*} \sim T(n-1)$$
$$\frac{\sqrt{n}\frac{\overline{X}-\mu}{\sigma}}{\sqrt{\frac{1}{n-1}\frac{(n-1)S^2}{\sigma^2}}} = \sqrt{n}\frac{\overline{X} - \mu}{S}$$
Доверительный интервал:

$\overline{X} - \frac{q_{1-\frac{\alpha}{2}}S}{\sqrt{n}}, \overline{X} + \frac{q_{1-\frac{\alpha}{2}}S}{\sqrt{n}}$

Определение. Распределение Фишера

$\Xi_n^2 \sim \Xi^2(n)$

$\Xi_m^2 \sim \Xi^2(m)$

Они независимы

$\frac{\frac{\Xi_n^2(n)}{n}}{\frac{\Xi_m^2(m)}{m}} \sim F(n,m)$

Мы доказали теорему Фишера давайте теперь с помощью теоремы мы рассмотрим задачу построения доверительных интервалов нормального закона $P(-q_{1-\frac{\alpha}{2}} \leq \frac{X-\mu}{\sigma^2} \leq q_{1-\frac{\alpha}{2}})$ - ну и потом просто выражаем

- σ^2 - известно, $\mu = ?$

  Рассмотрим два варианта: $\frac{X_1 - \mu}{\sigma} \sim N(0,1)$

  $\sqrt{n}\frac{\overline{X} - \mu}{\sigma} \sim N(0,1)$

  Доверительный интервал уровня $1-\alpha$ $[\overline{X} - \frac{q_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}, \overline{X} + \frac{q_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}]$

- μ - известно, $\sigma^2 = ?$

  $-q \leq \sqrt{n}\frac{(\overline{X} - \mu)}{\sigma} \leq q$

  $-q\sigma \leq \sqrt{n}(\overline{X} - \mu) \leq q\sigma$

  $\frac{\sqrt{n}(\overline{X} - \mu)}{q} \leq \sigma$

  $-\frac{\sqrt{n}(\overline{X} - \mu)}{q} \leq \sigma$

Рассмотрим следующую статистику:

$\sum \frac{(X_i - \mu)^2}{\sigma^2} \sim \Xi^2(n)$

$P(q_{\frac{\alpha}{2}} \leq \sum \frac{(X_i - \mu)^2}{\sigma^2} \leq q_{1-\frac{\alpha}{2}}) = 1 - \alpha$

Доверительный интервал:

$\frac{\sum (X_i - \mu)^2}{q_{1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{\sum (X_i - \mu)^2}{q_{\frac{\alpha}{2}}}$

Давайте теперь рассмотрим задачу построения доверительного интревала $\mu = ?, \sigma^2 = ?$

Воспользуемся теоремой Фишера:

$\frac{nS_*^2}{\sigma^2} \sim \Xi^2(n-1)$

Доверительный интервал: $\frac{nS_*^2}{q_{1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{nS_*^2}{q_{\frac{\alpha}{2}}}$

20 Асимптотические доверительные интервалы. Доверительный интервал для математического ожидания, дисперсии, медианы.

Раньше мы говорили $P(\theta \in (l_n, r_n)) \geq 1 - \alpha$
Таперь же мы будем говорить $\lim\limits_{n\to\infty} P(\theta \in (l_n, r_n)) \geq 1 - \alpha$

$T(X, \theta) \xrightarrow{d} G$ не зависит от θ
ЦПТ и ее следствия
$\sqrt{n}\dfrac{\overline{X} - \mu}{\sigma} \to N(0, 1)$
$\sqrt{n}\dfrac{\overline{X} - \mu}{S} \to N(0, 1)$
Доверительный интервал: $\overline{X} \pm \dfrac{q_{1-\frac{\alpha}{2}} S}{\sqrt{n}}$
$\sqrt{n}\dfrac{S_*^2 - \sigma^2}{\sqrt{\widehat{\beta_4} - S_*^4}} \to N(0, 1)$

Доверительный интервал: $S_*^2 \pm \dfrac{q_{1-\frac{\alpha}{2}} \sqrt{\widehat{\beta_4} - S_*^4}}{\sqrt{n}}$
Теорема об ассимтотике среднего члена вариационного ряда
$\sqrt{n}\dfrac{X_{(\lfloor np \rfloor)} - q_p}{\sqrt{p(1-p)}} f(q_p) \to N(0, 1)$
Доверительный интервал для медианы: $p = \dfrac{1}{2}$
$\sqrt{n} f(q_p) \dfrac{X_{(\lfloor \frac{n}{2} \rfloor)} - q_p}{\frac{1}{2}}$ (зачастую f это константа)
Доверительный интервал: $X_{(\lfloor \frac{n}{2} \rfloor)} \pm \dfrac{q_{1-\frac{\alpha}{2}}}{\sqrt{n} \cdot const}$

21 Постановка задачи проверки статистических гипотез: выбор нулевой и альтернативной гипотез, общий принцип работы статистического теста, p-value, ошибки I и II рода (false positive and false negative).

Нам надо будет выделить основное предположение (по умолчанию) и альтернативное предположение (наше подозрение или то, что мы хотим доказать)
Давайте рассуждать:
рациональное, с точки зрения инопланетянина (не опираться на жизненный опыт)

$X_1, \ldots, X_n$ - выборка в широком смысле.
$(X_1, \ldots, X_n) \sim F$
H_0 - нулевая гипотеза.
H_1 - альтернатива.
Так же пусть нам дали уровень значимости
$\alpha \in (0, 1)$ (по умолчанию 0.1, 0.05, 0.01, 0.001)
Статистический тест (критерий)
$\delta(X, \alpha, H_0, H_1) = \begin{cases} accept H_0 \\ reject H_0 (w.\ respect \text{ to } H_1) \end{cases}$ То есть в первом случае данные не противоречат H_0, а втором противоречат.
Но это не значит, что мы доказали утверждение.
Пусть у нас есть функция $T(X)$ - статистика критерия
$T(X)$ либо в точности, либо в пределе стремится к G при условии $H_0 (\sim$ or $\to)$
$P(T(X) \in T_0(\alpha) | H_0) = 1 - \alpha$
if $T(x) \in T_1(\alpha)$: reject H_0
else: acceptH_0
$T_0(\alpha)$ - область принятия
$T_1(\alpha)$ - область опровержения

1. left: $T_0(\alpha) = [q_\alpha, +\infty)$
2. right: $T_0(\alpha) = (-\infty, q_\alpha]$
3. two: $T_0(\alpha) = [q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}], T_1(\alpha) = \overline{T_0(\alpha)}$

$p_l = P(U \leq T(x) | H_0)$
$p_r = P(U > T(x) | H_0)$
$p = 2min(p_l, p_r)$
if p $< \alpha$: reject H_0
else: accept H_1
Если мы опровергли нулевую гипотезу, но она верна, то это ошибка первого рода (false positive), ее максимальная вероятность это в точности α
Если мы не опровергли нулевую гипотезу, но она была не верна, то это ошибка второго рода (false negative), ее вероятность это β
$\beta = P(T(X) \in T_0(\alpha) | H_1)$

22 Статистические тесты, основанные на доверительных интервалах (z-тест для одной-/двух выборок, t-тест для одной/двух выборок, F-тест).

Когда мы строили доверительные интервалы то мы зажимали статистику между квантилями. Это похоже на двухсторонний тест.
$X_1, \dots, X_n \sim F_\theta$
$T(X, \theta) \to U \sim G$
$P(q_{\frac{\alpha}{2}} \le T(X, \theta) \le q_{1-\frac{\alpha}{2}}) = 1 - \alpha$
$H_0 : \theta = \theta_0$
$P(T(X, \theta_0) \in T_0(\alpha) | \theta = \theta_0) = 1 - \alpha$
$H_1 = \theta \ne \theta_0, \theta > \theta_0, \theta < \theta_0$
Пример:
1) $X_1, \dots, X_n \sim F, \mu = EX_1, \exists \operatorname{Var} X$
$H_0 : \mu = \mu_0$
$T(X) = \sqrt{n}\dfrac{\overline{X} - \mu_0}{S} \to N(0,1)$ если $\mu = \mu_0$
При $H_1 : \mu \ne \mu_0$ у нас двухсторонняя критическая область
При $H_1 : \mu > \mu_0$ у нас правосторонняя критическая область
При $H_1 : \mu < \mu_0$ у нас левосторонняя критическая область
Дальше для этого приводится пример с больницей (нам либо надо просто проверить, что температура не стандартная, либо нам важно знать, что она больше нормы, либо нам важно знать, что она ниже нормы)
$P(\sqrt{n}\dfrac{\overline{X} - \mu_0}{S} \in T_0(\alpha) | \mu \ne \mu_0) = P(\sqrt{n}\dfrac{\overline{X} - \mu}{S} + \dfrac{\mu - \mu_0}{S}\sqrt{n} | \mu \ne \mu_0)$
Это будет стремиться либо к $\Phi(-\infty)$ либо $1 - \Phi(+\infty)$

23 Критерии Колмогорова-Смирнова.

Критерий Колмогорова
$X_1, \dots, X_n \sim F$
$H_0 : F = F_0$ (F_0 - непр)
$H_1 : F \ne F_0$
Идея основана на теореме Колмогорова (было в начале семестра)
$D_n = \sqrt{n} \underset{x \in \mathbb{R}}{sum} |F_n(x) - F_0(x)|$, F_n - эмпирическая функция распределения
if $D_n > q_{1-\alpha}$ then reject H_0 else accept H_0
1) $n \ge 20$ работает хорошо, при маленьких n есть спец таблицы
2) Так же есть приближенные формулы для D_n
3) $H_0 : F = F(\theta), H_1 : \neg H_0 \Rightarrow D_n = \sqrt{n} \underset{x}{sup} |F_n(x) - F_0(x, \theta)|$
$\theta \to \widehat{\theta}$
В пределе будет более сложная формула
Критерий Смирнова

$X_1, \dots, X_n$
$Y_1, \dots, Y_m$
Они независимы
$H_0 : F_X = F_Y (= F_0)$
$H_1 :\neq H_0$
Тут идея основана на формуле Смирнова
$D_{n,m} = \sqrt{\frac{nm}{n+m}} sup_x |F_n(x) - F_m(x)|$
$T_1(\alpha) = (q_{1-\alpha}, +\infty)$

24 Критерий согласия Пирсона хи-квадрат для простой и сложных гипотез

$X_1, \dots, X_n \sim F(x)$ - непрерывная
Давайте дискритезируем данные
$\Delta_1 : \nu_1$ - количесво элементов выборки попадающих Δ_1
...
Δ_N
$p_{\Delta_k} = \int_{\delta_k} p(x)dx$, $p(x) = F'(x)$
Рассмотрим $\{1, 2, \dots, N\}$, $p = (p_1, \dots, p_N)$ - настоящий вектор вероятностей
$p_0 = (p_{01}, \dots, p_{0N})$ - ожидаемый фиксированный вектор вероятностей
ν_k - количество элементов в выборке типа k
$H_0 : p = p_0$
$H_1 : p \neq p_0$
$n = \sum_{k=1}^{N} \nu_k$
$\Xi_N^2 = \sum_{k=1}^{N} \frac{(\nu_k - np_{0k})^2}{np_{0k}}$
Теорема
$\Xi_N^2 \xrightarrow[n\to\infty]{} \Xi^2(N-1)$ при условии H_0
Доказательство
$N = 2$:
$\frac{(\nu_1 - np_{01})^2}{np_{01}} + \frac{(\nu_2 - np_{02})^2}{np_{02}} = \frac{(\nu_1 - np_{01})^2}{np_{01}} + \frac{(n - \nu_1 - n(1-p_{01}))^2}{n(1-p_{01})} = \frac{(\nu_1 - np_{01})^2}{n}(\frac{1}{p_{01}} + \frac{1}{1-p_{01}}) = \frac{(\nu_1 - np_{01})^2}{np_{01}(1-p_{01})}$
Без квадрата по ЦПТ это стремится к N(0, 1), но мы можем навесить непрерывную функцию возведения в квадрат и получим то, что нам нужно
Для классического критерий хи-квадрат у нас правосторонняя критическая область (потому что в сумме при H_0 будет маленькая разность в квадрате и тд и тп очев крч)
Критерий состоятельный (то есть вероятность ошибки второго рода стре-

мистя к единице, если погнать объем выборки на бесконечность)
Теорема
$p_0(\theta) > 0 \forall \theta$
$\frac{\partial p_0}{\partial \theta}, \frac{\partial^2 p_0}{(\partial \theta)^2}$ - непрерывная
Тогда $\Xi_N^2 \to X^2(N-1-d)$
$rk(\frac{\partial p_{0k}}{\partial \theta_j})_{1 \le k \le N, 1 \le j \le d} = d$

25 Критерий однородности хи-квадрат, в том числе случай 2 × 2

Пусть у нас k независимых выборок все они из $\{1, 2, \dots, N\}$
Пусть $p^{(j)}$ - истинный вектор вероятностей для соответствующей выборки
$H_0 : p^{(1)} = \dots = p^{(k)}$
$H_1 :\neq H_0$
ν_{ij} - количество элементов типа j в i-ой выборке
$n_i = \sum\limits_j \nu_{ij} = \nu_{i*}$ - объем i-ой выборки
$n = n_1 + \dots + n_k$
Пусть $p^{(1)}, \dots, p^{(k)}$ - известны
$\Xi_{n_1}^2 = \sum\limits_{j=1}^{N} \frac{(\nu_{ij} - n_i p_j^{(i)})^2}{n_i p_j^(i)}, df = N - 1$
$\Xi_{n_1, n_2, \dots, n_k}^2 = \sum\limits_{i=1}^{K} \Xi_{n_i}^2, df = k(N-1)$
Рассмотрим $p^{(1)}, \dots, p^{(k)}$ - не известны
$df = k(N-1) - (N-1) = (N-1)(k-1)$
$\widehat{p_j} = \frac{\nu_{1j} + \dots + \nu_{kj}}{n} = \frac{\nu_{*j}}{n}$
$L(\dots) = p_1^{(\nu_{*1})} \dots p_k^{(\nu_{*k})}, n = \sum\limits_j \nu_{*j}$
$Z_{n_1 n_2} = (\frac{\nu_{11}}{n_1} - \frac{\nu_{21}}{n_2}) \sqrt{\frac{n n_1 n_2}{\nu_{*1} \nu_{*2}}}$
Утверждается, что $Z_{n_1 n_2}^2 = \Xi_{n_1 n_2}^2 \to N(0,1)$
$H_0 : p_1 = p_2, p_j = P(0$ в j-ой выборке$)$
H_1 - настраивается

26 Критерий независимости хи-квадрат, в том числе случай 2 × 2

$X_1, \dots, X_n : \{1, 2, \dots, N\}$
$Y_1, \dots, Y_n : \{1, 2, \dots, M\}$

ν_{ij} - количество пар, в которых первая компонента равна i, а вторая j
Это можно представить в виде таблицы сопряженности
Проссумируем по каждому столбцу и по каждой строчке
Пусть $p_{ij} = P(X = i, Y = j)$
$p_{xi} = P(X = i)$
$p_{yj} = P(Y = j)$
$H_0 : p_{ij} = p_{xi}p_{yj} \forall i, j$
$H_1 : \neg H_0$
$\Xi^2 = \sum\limits_{1 \le i \le N, 1 \le j \le M} \frac{(\nu_{ij} - p_{ij}n)^2}{np_{ij}}, df = MN - 1$
$df = MN - 1 - (N - 1) - (M - 1) = MN - N - M + 1 = N(M - 1) - (M - 1) = (M - 1)(N - 1)$
$\widehat{p_{Xi}} = \frac{\nu_{i*}}{n}$
$\widehat{p_{Yj}} = \frac{\nu_{j*}}{n}$
$Z_n = (\frac{\nu_{11}}{\nu_{1*}} - \frac{\nu_{21}}{\nu_{2*}})\sqrt{\frac{n\nu_{1*}\nu_{2*}}{\nu_{*1}\nu_{*2}}}$
Утверждается, что $Z_n^2 = \Xi^2$
$Z_n = \sqrt{n}\rho_n$, ρ_n - выборочный коэффициент корреляции Пирсона (ковариация делить на корень из произведения дисперсий ыыыыыыыы)
$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{E(XY) - EXEY}{\sigma_X\sigma_Y} = \frac{P(X=1,Y=1) - P(X=1)P(Y=1)}{\sqrt{P(X=1)P(X=0)P(Y=1)P(Y=0)}} =$
... Тренер ну тут очев лол кек кд чд бро $= P(Y = 1|X = 1) - P(Y = 1|X = 0)$

27 Критерий квантилей и знаков

27.1 Критерий квантилей

H_0 :
$F(q_1) = \alpha_1$
$F(q_2) = \alpha_2$
...
$F(q_N) = \alpha_N$
где $\alpha_0 = 0 < \alpha_1 < \cdots < \alpha_N < 1 = \alpha_{N+1}$
$H_1 : \neg H_0$
$q_0 = \inf \mathrm{supp}\, P < q_1 < \cdots < q_N < \sup \mathrm{supp}\, P = q_{N+1}$
$\Delta_1 = [q_0, q_1)$
...
$\Delta_{N+1} = [q_N, q_{N+1})$
$P(\Delta_1) = \alpha_1 - \alpha_0$
...
$P(\Delta_N) = \alpha_{N+1} - \alpha_N$

27.2 Критерий знаков

Это особый случай $H_0 : F(q) = \frac{1}{2}$ $(X_1, Y_1)^T, \dots, (X_n, , Y_n)^T$
Мы хотим проверить что:
а) выборки независимы
б) распределения одинаковы
$F(x, y) = F_1(x) \cdot F_1(y)$
$U = X - Y \Rightarrow \operatorname{med} U = 0$
ν_1 - количество элементов новой выборки $>$ med
$Z_n = \frac{2}{\sqrt{n}}(\nu_1 - \frac{n}{2}) \to N(0, 1)$
$Z_n = \sqrt{n}\rho_n, \rho_n$ - коэффициент корреляции Пирсона
Теорема
$(X_1, Y_1)^T, \dots, (X_n, Y_n)^T \sim N(\dots, \dots) \Rightarrow \frac{\sqrt{n-2}\rho_n}{\sqrt{1-\rho_n^2}} \sim T(n-2)$
$H_0 : \rho = 0$
$H_1 : (\rho \neq 0, \rho > 0, \rho < 0)$

28 Ранговые критерии. Критерии Манна-Уинтни-Уилкоксона

Определение. Ранг
$X_1, \dots, X_n$ - выборка
$r(X_k)$ - номер X_k в вариационном ряде
$X_1, \dots, X_n$
$Y_1, \dots, Y_m$
Это две независиммых выборки, давайте объединим их в одну
Рассмотрим $(X_1, \dots, X_n, Y_1, \dots, Y_m)$
R_i - ранг X_i, в объединенной выборке
$T = R_1 + \dots, R_n$ - Статистика Вилкоксона
$Z_{rs} = \mathbb{1}(X_r < Y_s)$
$U = \sum_{r=1}^{n} \sum_{s=1}^{m} Z_{rs}$ - Мант-Уитни
$T + U = mn + \frac{n(n+1)}{2}$
Хотим проверить, что распределение X совпадает с Y
$EU = mnE\mathbb{1}(X < Y) = mnP(X < Y) = \frac{1}{2}$ - при условии H_0
$H_0 : P(X < Y) = \frac{1}{2}$
$U \sim N(\frac{mn}{2}, \frac{nm(m+n+1)}{12})$
Теорема
$(X_1, Y_1)^T, \dots, (X_n, , Y_n)^T \sim N(\dots, \dots) \Rightarrow \frac{\sqrt{n-2}\rho_n}{\sqrt{1-\rho_n^2}} \sim T(n-2)$

29 Коэффициенты корреляции Пирсона, Спирмена и Кендала. Статистические тесты, основанные на них.

Хотим проверить независимость
R_i - ранг X_i (в своей выборке)
S_i - ранг Y_i (в своей выборке)
ρ - выборочный коэффициент корреляции между R_i и S_i - коэффициент корреляции Спирмена
$\rho = \frac{12}{n(n^2-1)} \sum (R_i - \frac{n+1}{2})(S_i - \frac{n+1}{2})$
$H_0 : \rho = 0, \sqrt{n}\rho \to N(0,1)$
$H_1 : \rho \neq 0, \rho > 0, \rho < 0$
Коэффициент корреляции Кендала
$\tau = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sign}(T_j - T_i)$
H_0 верно $\Rightarrow E\tau = 0, \operatorname{Var}\tau = \frac{2(2n+5)}{9n(n-1)}$
$\tau \approx N(0, \frac{4}{9n})$
$H_0 : \tau = 0$
$H_1 :> 0, < 0, \neq 0$
$\rho = \frac{\operatorname{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y}$
Пирсон проверяет линейную зависимость между двумя случайными величинами
Спирмен проверяет монотонную зависимость между двумя случайными величинами
Кендал делает тоже самое что и Спирмен

30 Критерий инверсий

$X_{(1)} \leq \cdots \leq X_{(n)}$ Крайние ситуации: выборка отсортированна, то есть трудно поверить, что у нас все случайно
ν_i - количество инверсий для элемента X_i
$T = \nu_1 + \cdots + \nu_n$
$ET = \frac{n(n-1)}{4}$
$\operatorname{Var} T = \frac{n(n-1)(2n+5)}{72}$

Статистика T - ассимптотически нормальная

31 Модель линейной регрессии. Минимальные и обычные предположения. Оценка наименьших квадратов.

$Y = Xb + \varepsilon$
$X \in M_{n\times m}(\mathbb{R})$ - матрица переменных
x_{ij} - количественная переменная
$Y \in \mathbb{R}^n$ - наблюдения зависимой переменной
$b \in \mathbb{R}^m$ - неизвестный вектор коэффициентов
$\varepsilon \in \mathbb{R}^n$ - случайная ошибка
1) $E\varepsilon = 0$
2) $\operatorname{Var}\varepsilon_i = \sigma^2$ - гомоскедотичность
3) $\operatorname{Cov}(\varepsilon_i, \varepsilon_j) = 0$
Наша цель оценить b и σ^2
Определение. Ошибка наименьших квадратов
$\widehat{b} = \operatorname{argmin} S^2(b)$
$S^2(b) = (Xb - Y)^T(Xb - Y)$
$A = X^TX \in M(m \times n)$
$\frac{1}{n}X^TX$
$rank(A) = m$
Теорема
$\widehat{b} = A^{-1}X^TY$
Доказательство
$S^2(\widehat{b} + \delta) = S^2(\widehat{b}) + \delta^T A\delta$
$t = Tb, T \in M_{k\times m}, k \le m, rankT = k$
$\widehat{t} = T\widehat{b}$

32 Теорема Гаусса-Маркова

1) $\widehat{t}$ - несмещенная оценка t
2) $\widehat{t}$ - оптимальная оценка в классе линейных несмещенных оценок
Доказательство
$E\widehat{t} = ET\widehat{b} = ETA^{-1}X^TY = TA^{-1}X^TEY = TA^{-1}X^TXb = Tb = t$
$\operatorname{Var}(\widehat{t}) = \operatorname{Var} TA^{-1}X^TY = TA^{-1}X^T \operatorname{Var} YXA^{-1}T^T = \sigma^2TA^{-1}T^T$
$\operatorname{Var}\widehat{t} = \sigma^2TA^{-1}T^T$
$\operatorname{Var}\widehat{b} = \sigma^2A^{-1}$
Пусть LY - несмещенная оценка для t, т.е.
$ELY = t = Tb \Rightarrow \operatorname{Var} LY = \sigma^2LL^T$
Заметим, что $LL^T = (TA^{-1}X^T)(TA^{-1}X^T)^T + (L - TA^{-1}X^T)(L - TA^{-1}X^T)^T$

$MSE\widehat{t} = tr\,\mathrm{Var}\,\widehat{t}$
Тогда $L = TA^{-1}X^T$

33 Оценка остаточной дисперсии

$ES^2(b) = E(Xb-Y)^T(Xb-Y) = E\varepsilon^T\varepsilon = n\sigma^2$
$E(\widehat{b}-b)A(\widehat{b}-b) = \sum\limits_{i,j} a_{ij}E(\widehat{b_i}-b_i)(\widehat{b_j}-b_j) = \sum\limits_{i,j} a_{ij}\,\mathrm{Cov}(\widehat{b_i},\widehat{b_j}) = \sigma^2\sum\limits_{i,j} a_{ij}a_{ij}^{-1} =$
$\sigma^2\sum\limits_i\sum\limits_j a_{ij}a_{ji}^{-1} = \sigma^2\sum\limits_{i=1}^{m} 1 = \sigma^2 m$
$S^2(b) = S^2(\widehat{b}) + (\widehat{b}-b)^TA(\widehat{b}-b) \Rightarrow n\sigma^2 = ES^2(\widehat{b}) + m\sigma^2 \Rightarrow \widehat{\sigma}^2 = \dfrac{S^2(\widehat{b})}{n-m}$ - несмещенная оценка для σ^2

34 Условная оценка наименьших квадратов

$Tb = t_0, T \in M_{k\times m}, rankT = k$
$\widehat{b}_{T,t_0} = \underset{Tb=t_0}{\mathrm{argmin}} S^2(b)$ - условная оценка наименьших квадратов
Теорема
$\widehat{b}_{T,t_0} = \widehat{b} - A^{-1}T^TD^{-1}(T\widehat{b}-t_0), D = TA^{-1}T^T$
Можно показать: $S^2(b) = S^2(\widehat{b}_{T,t_0}) + (\widehat{b}_{T,t_0}-b)^TA(\widehat{b}_{T,t_0}-b)$

35 Основная теорема о линейной регрессии. Следствия. t-тест.

1) $S^2(\widehat{b}), \widehat{b}$ - незав.
$S^2(b) - S^2(\widehat{b}), S^2(\widehat{b})$ - незав.
2) $\widehat{b} \sim N(b, \sigma^2A^{-1})$
3) $\dfrac{S^2(\widehat{b})}{\sigma^2} \sim \Xi(n-m)$
$\dfrac{S^2(b)-S^2(\widehat{b})}{\sigma^2} \sim \Xi^2(m)$
Без доказательства
T-test:
$$\frac{\frac{\widehat{b_j}-b_j}{\sigma\sqrt{A_{ij}^{-1}}}}{\sqrt{\frac{1}{n-m}\frac{S^2(\widehat{b})}{\sigma^2}}} \sim T(n-m)$$

$$\sqrt{\frac{1}{n-m}\frac{S^2(\widehat{b})}{\sigma^2}} = \frac{\widehat{b_j} - b_j}{S(\widehat{b}\sqrt{A_{ij}^{-1}})}\sqrt{n-m}$$
$H_0 : b_j = b_{0j}$
$H_1 : <, \neq, >$

36 F-тест. Коэффициент детерминации.

$T \in M_{k\times m}$
$H_0 : Tb = t_0$ (default: T = E_m, $t_0 = 0$, то есть все $b_i = 0$ одновременно)
$H_1 : Tb \neq t_0$
$$F = \frac{n-m}{k}\cdot\frac{S^2(\widehat{b}_{T,t_0}) - S^2(\widehat{b})}{S^2(\widehat{b})} = \frac{n-m}{k}\cdot\frac{(T\widehat{b} - t_0)^T D^{-1}(T\widehat{b} - t_0)}{S^2(\widehat{b})} \sim F(k, n-m)$$
$$R = \frac{\sum_i (Y_i - \overline{Y})(\widehat{Y}_i - \overline{\widehat{Y}})}{\sqrt{\sum_i (Y_i - \overline{Y})^2 \sum_i (\widehat{Y}_i - \overline{\widehat{Y}})^2}},\ \widehat{Y} = X\widehat{b}$$
R - многомерный коэффициент корреляции
R^2 - коэффициент детерминации

37 Однофакторный дисперсионный анализ

$Y_{ij} = \mu_j + \varepsilon_{ij}$, μ_j - среднее влияние на $j-$ом факторе
$1 \leq j \leq \nu$ - уровень фактора
$1 \leq i \leq I_j$ - i-ое наблюдение для фатора с уровнем j
$$I = \sum_{j=1}^{\nu} I_j$$
$\varepsilon_{ij} \sim N(0, \sigma^2 E)$
$H_0 : \mu_1 = \mu_2 = \cdots = \mu_J$
$H_1 : \neg H_0$
$$F = \frac{\frac{S_B^2}{df_B}}{\frac{S_W^2}{df_W}} \sim F(J-1, I-J)$$
$$S_B^2 = \sum_{j=1}^{\nu} I_j(\overline{Y_{*j}} - \overline{Y})^2$$ - межгруповая дисперсия
$df_B = J - 1$
$$S_W^2 = \sum_i \sum_j (Y_{ij} - \overline{Y_{*j}})^2$$ - внутригруповая дисперсия
$df_W = I - J$
$S^2 = S_W^2 + S_B^2$

38 Двухфакторный дисперсионный анализ.

$Y_{ij} = \mu + a_i + b_j + \varepsilon_{ij}$
$1 \le i \le I$
$1 \le j \le J$
Для каждой пары (i, j) ровно одно наблюдение
M - общее среднее
a_i - среднеее влияние на i-ом уровне фактора 1
b_j - среднеее влияние на j-ом уровне фактора 2
Факторы независимы
$\sum a_i = \sum b_j = 0$
ε_{ij} - незав.
$\varepsilon_{ij} \sim N(0, \sigma^2)$

$$H_A : \forall a_i = 0 \to F_A = \frac{(I-1)(J-1)}{I-1} \frac{S_A^2}{S_{ALL}^2}$$

$$H_B : \forall b_i = 0 \to F_B = \frac{(I-1)(J-1)}{J-1} \frac{S_B^2}{S_{ALL}^2}$$

$$H : \forall a_i = b_j = 0 \to F = \frac{(I-1)(J-1)}{I+J-2} \frac{S_A^2 + S_B^2}{S_{ALL}^2}$$

$$S_A^2 = J \sum_i (\overline{Y_{i*}} - \overline{Y})^2$$

$$S_B^2 = I \sum_j (\overline{Y_{*j}} - \overline{Y})^2$$

$$S_{ALL}^2 = \sum_{i,j} (Y_{ij} - \overline{Y_{i*}} - \overline{Y_{*j}} + \overline{Y})^2$$

39 Ковариационный анализ.

1 фактор, 1 количественная переменная
$y_{ij} = \beta_i + \gamma z_{ij} + \varepsilon_{ij}$
$1 \le i \le I$
$1 \le j \le J$
$n = I \cdot J$
На каждой паре (i, j) одно наблюдение
β_i - среднее влияние на уровне i
z_{ij} - количественная переменная
ε_{ij} - нез.
$\varepsilon_{ij} \sim N(0, \sigma^2)$
$H_\gamma : \gamma = 0$

$$\widehat{\beta}, \widehat{\gamma} \text{ - ОНК} \Rightarrow \sqrt{n-I-1} \frac{\widehat{\gamma}}{\sqrt{A_{\gamma\gamma}^{-1} S^2(\widehat{\beta}, \widehat{\gamma})}} \sim T(n-I-1)$$

$A_{\gamma\gamma} = A_{(I+1)(I+1)}$
$S^2(\widehat{\beta}, \widehat{\gamma}) = \sum_{ij} (y_{ij} - \widehat{\beta}_i - \widehat{\gamma} z_{ij})^2$ - суммарная квадратичная ошибка

$H_B : \beta_1 = \cdots = \beta_I$

$$F = \frac{n-I-1}{I-1} \cdot \frac{\sum\limits_{i,j} y_{ij}^2 - n\overline{y}^2 - p}{p}$$

$$p = \sum_{i,j} y_{ij}^2 - n\overline{y} - \frac{(\sum\limits_{i,j} z_{ij}y_{ij} - n\overline{z})^2}{\sum\limits_{i,j} z_{ij}^2 - n\overline{z}}$$

40 Логистическая регрессия

$Y_i \approx \frac{1}{1+e^{-x_i c}}$, $X = (X_1 \dots X_n)^T$ $c = (c_1 \dots c_m)^T$

$Y_i \sim Bern(\frac{1}{1+e^{-x_i c}})$

$L(X,c) = \prod\limits_{i=1}^{n} (\frac{1}{1+e^{-x_i c}})^{Y_i} (1 - \frac{1}{1+e^{-x_i c}})^{1-Y_i}$

$-\ln L(X,c) = \dots$

41 Последовательный анализ Вальда. Описание алгоритма. Теорема о конечном числе шагов для простых гипотез.

$H_0 : F = F_0$

$H_1 : F = F_1$ f_0, f_1 - соответствующие плотности $A_0 < 1 < A_1$

$n = 1$

while true {

if $\frac{L_{0n}}{L_{1n}} \le A_0$ then return reject H_0

if $\frac{L_{0n}}{L_{1n}} \ge A_1$ then return accept H_0

$++n$ }

$\alpha, \beta, a_0 = \ln A_0, a_1 = \ln A_1$

$Z_i = \ln \frac{f_1(X_i)}{f_0(X_i)}$

$$\sum_{i=1}^{n} Z_i = \ln \frac{L_{1n}}{L_{0n}}$$

Теорема

Пусть ν - количество итераций

$\Rightarrow P(\nu \ge n | H_0) \to 0, P(\nu \ge n | H_1 \to 0), n \to \infty$

Доказательство

$r - fixed$

$\eta_1 = Z_1 + \cdots + Z_r$
$\eta_1 = Z_{r+1} + \cdots + Z_{2r}$
$\ldots$ $\{\nu \geq rk\} \Leftrightarrow \{a_0 < \eta_1 + \cdots + \eta_j < a_1\}_{i \leq j \leq k} = \{a_0 < \eta_1 < a_1, a_0 < \eta_1 + \eta_2 < a_1, \ldots, a_0 < \eta_1 + \cdots + \eta_k < a_1\} \subset \{|\eta_j| < b = a_i - a_0\}_{1 \leq j \leq k} \Rightarrow P(\nu \geq rk|H_s) \leq P_{1 \leq j \leq k}(|\eta_j| < b|H_s) = P_{1 \leq j \leq k}(\nu_j^2 < b^2|H_s) = P^k(\eta_1^2 < b^2|H_s) = p_s^k$
$E_s\eta_1^2 \geq \operatorname{Var}_s \eta^1 = r \operatorname{Var}_s Z_1 > b^2, r > \max\left(\frac{b^2}{\operatorname{Var}(Z_1|H_0)}, \frac{b^2}{\operatorname{Var} Z_1|H_1}\right)^{\substack{E_s(\ldots)=E(\ldots|H_s) \\ \operatorname{Var}_s(\ldots)=\operatorname{Var}(\ldots|H_s)}} \Rightarrow$
$p + s < 1$
$P(\nu > n|H_s) \leq P(\nu \geq rk|H_s) \leq p_s^k \to 0, k = k(n), n \to \infty$

42 Последовательный анализ Вальда. Теорема об оценке граничных констант и вероятностей ошибок первого и второго рода. Оценка среднего числа итераций.

Теорема

Пусть (α, β) - вероятности $\Rightarrow A_0 \geq A_0' = \frac{\beta}{1-\alpha}$ и $A_1 \leq A_1' = \frac{1-\beta}{\alpha}$
if A_0' и $A_1 = A_1'$ then α', β' - вероятности ошибок
$\alpha' \leq \frac{\alpha}{1-\beta}, \beta' \leq \frac{\beta}{1-\alpha}$
$\alpha' + \beta' \leq \alpha + \beta$
Доказательство
$1 = \sum_{n=1}^{\infty} P(\nu = n|H_0) = \sum_{n=1}^{\infty} P(\text{accept at the step n} \mid H_0) + \sum_{n=1}^{\infty} P(\text{reject at the step n} \mid H_0) =$
$\alpha = \sum_{n=1}^{\infty} P(\text{reject at the step n} \mid H_0) \leq \frac{1}{A_1} \sum_{n=1}^{\infty} P(\text{reject at the step n} \mid H_1)$
$L_{n1} \geq L_{n0} \cdot A_1$
$L_{n1 \geq L_{n0} \cdot A_1}$
Аналогичным образом мы можем получить неравенство:
$\beta \leq A_0(1-\alpha)$
$\alpha \leq \frac{1}{A_1}(1-\beta)$
Пусть $A_0' = \frac{\beta}{1-\alpha}$ и $A_1' = \frac{1-\beta}{\alpha}$ и ошибки $(\alpha', \beta') \Rightarrow \frac{\beta}{1-\alpha} \geq \frac{\beta'}{1-\alpha'}$
$\frac{1-\beta}{\alpha} \leq \frac{1-\beta'}{\alpha'} \Rightarrow \alpha' \leq \frac{\alpha(1-\beta')}{1-\beta} \leq \frac{\alpha}{1-\beta}$
$\beta(1-\alpha) \leq \beta(1-\alpha')$
$(1-\beta)\alpha' \leq (1-\beta')\alpha$
Сложим последние два неравства и получим требуемое
Давайте вспомним тождество Вальда:
(X_i) - независимые случайные величины с математическим ожиданием a

ν - целочисленная случайная величина не зависящая от X_i
$E\nu = b$
$S_\nu = X_1 + \cdots + X_\nu$
$\Rightarrow ES_\nu = ab = E(ES_\nu|\nu) = E\nu a = aE\nu = ab$
$a_0 = \ln \dfrac{\beta}{1-\alpha}, a_1 = \ln \dfrac{1-\beta}{\alpha}$
$E_s\nu \cdot E_s Z_i = E_s S_\nu \approx a_0(1-\alpha) + a_1(1-\beta)$
$E_1\nu \cdot E_1 Z_i \approx \beta a_0 + a_1(1-\alpha)$

43 Бутстреп. Оценка дисперсии, построение д.и., permutation test.

1. $X_1, \ldots, X_n \sim F_\theta; \widehat{\theta}; \operatorname{Var}\widehat{\theta}-?$ Можно ли ее оценить численно?
2. $X_1, \ldots, X_n; \overline{X}; \operatorname{Var}\overline{X} = \dfrac{\sigma}{\eta}$ можно ли оценить численно?
3. $X_1, \ldots, X_n; \operatorname{med} X; \operatorname{Var}\operatorname{med}(X)-?$

$X-1, \ldots, X_n; \widehat{\theta}$ - оценка чего-то; F_n - Э.Ф.П
for (int i = 0; i <= B - 1; ++i) {
$X_1^*, \ldots, X_n^* \sim F_n$
$\theta_i^* = g(X_1^*, \ldots, X_n^*)$
$\theta^*.append(\theta^*)$
}
$\operatorname{Var}^* = \dfrac{1}{B}\sum_{i=0}^{B-1}(\theta_I^* - \overline{\theta_i^*})^2$
$\dfrac{\operatorname{Var}^*}{\operatorname{Var}\widehat{\theta}} \xrightarrow[n\to\infty]{P} 1$

43.1 CI

$\widehat{F(t)} = \dfrac{1}{B}\sum_{j=1}^{B}\mathbb{1}(\sqrt{n}(\overline{\theta_i^*} - \widehat{\theta}) \le t)$
$\widehat{\theta} \pm \dfrac{q_1 - \frac{\alpha}{2}}{\sqrt{n}}; q_{1-\frac{\alpha}{2}}$ - квантиль $\widehat{F}$
$\widehat{\theta} - \dfrac{q_{\frac{\theta}{2}}}{\sqrt{n}}; \widehat{\theta} + \dfrac{q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$

43.2 The permutation test

$X_1, \ldots, X_n \sim F$
$Y_1, \ldots, Y_m \sim G$
Both in dependent

$H_0 : F = G$
$H_1 : F \neq G$
$W = (X_1, \dots, X_n, Y_1, \dots, Y_m)$
$X_1 \dots X_n Y_1 \dots, Y_m$
$W_1 \dots W_{n+m}$
Пусть $T = T(X, Y)$ - некая статистика
$T = |\overline{X} - \overline{Y}|$
$T(W, Z) = |\overline{(Z, W)} - \overline{Y(Z, W)}|$
$X(Z, W) = \{W_i : Z_i = 1\}$
$Y(Z, W) = \{W_i : Z_i = 2\}$
Пусть Z^* - перестановка Z
for i in range(B):
Z^* - перестановка
$T^* = T(Z^*, W)$
$T_{all}.append(T^*)$
$p = \frac{1}{B} \sum_{j=1}^{B} \mathbb{1}(T_{all}[j] > t), t = T(W, Z)$
$p = p_{value}$

44 Введение в байесовскую статистику. Credible intervals. Проверка гипотез.

44.1 Credible interval

$P(a_l < \theta < a_r | X) = 1 - \alpha$
$\theta \sim \Gamma(\theta)$
$$p(\theta|X) = \frac{p(X|\theta) \cdot \pi(\theta)}{\int p(X|\theta)\pi(\theta)d\theta}$$

44.2 Hypothesis. Bayesian

$H_0 : \theta = \theta_0 - p_0 - prior$
$H_1 : \theta = \theta_1 - p_1 - prior$
$$L(\theta_0|X) = \frac{L(X|\theta_0) \cdot p_0}{const}$$
$const = P(X) = L(X|\theta_0) \cdot p_0 + L(X|\theta_1) \cdot p_1$
$L(\theta_1|X) \approx L(X|\theta_1) \cdot p_1$
$Y = Xc + \varepsilon, C \sim N(.,.)$
$\varepsilon \sim N(0, \sigma^2 I)$
$(c, \varepsilon) \sim (.,.)$
$(Y|C) \sim N(Xc, \sigma^2 I) \Rightarrow L(c|Y) \sim N(.,.)$
$\widehat{c} = E(c|Y)$