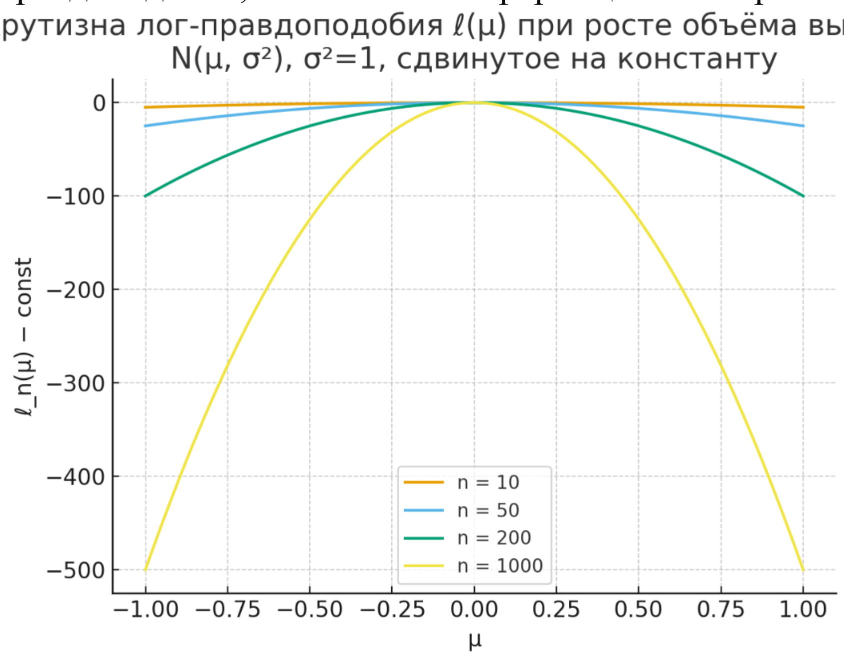
Эффективная оценка

Теорема (неравенство Рао-Крамера)

Пусть $X_1, X_2, \ldots X_n$ – выборка из параметрической модели с плотностью $p(x,\theta)$, пусть $\int p(\bar{x},\theta)\,d\bar{x}$ допускает дифференцирование по параметру. Тогда каждая несмещенная оценка $\hat{\theta}_{\boldsymbol{n}} = \varphi(X_1, X_2, \ldots X_n)$ параметра θ удовлетворяет неравенству:

$D\hat{\theta}_n \geq \dfrac{1}{nI_1(\theta)}$, где $I_1(\theta) = M\left(\dfrac{\partial \ln p(X_k,\theta)}{\partial\theta}\right)^2 = \int\limits_{-\infty}^{\infty}\left(\dfrac{\partial \ln p(x,\theta)}{\partial\theta}\right)^2 p(x,\theta)\,dx$ – информация Фишера относительно семейства $p(x,\theta)$, содержащаяся в одном наблюдении.

Информация Фишера — это число, которое показывает, насколько "чувствительны" ваши данные к изменению оцениваемого параметра.
Высокая информация Фишера: Если небольшое изменение параметра θ приводит к большому изменению в вероятности увидеть наши данные, значит, данные несут много информации о параметре. На основе таких данных можно построить очень точную (эффективную) оценку.
Представьте, что вы по звуку двигателя пытаетесь диагностировать неисправность. Если неисправность приводит к громкому и уникальному стуку (высокая чувствительность), то вы легко и точно поставите диагноз.
Низкая информация Фишера: Если данные выглядят практически одинаково при разных значениях параметра, то информации в них мало. Оценка будет иметь большой разброс.
Пример: Если все неисправности двигателя вызывают лишь тихий и похожий гул, вам будет очень сложно установить точную причину.
Формально: Информация Фишера I(θ) — это математическое ожидание квадрата производной логарифма функции правдоподобия. Чем "острее" функция правдоподобия, тем больше информация Фишера.



На графике видно: по мере роста n парабола становится более «узкой и острой». Это и есть рост кривизны/крутизны лог-правдоподобия, который пропорционален n. Следствие: дисперсия ММП-оценки падает как 1/n, а стандартная ошибка — как $\frac{1}{\sqrt{n}}$.

Замечание. Для дискретной модели вместо $p(x,\theta)$ используем $P(X = x_k)$.

Определение. Оценка $\hat{\theta}_n$ называется эффективной, если для нее неравенство Рао-Крамера превращается в равенство, то есть ее эффективность

$$e\left(\hat{\theta}_n\right) = \frac{1}{nI_1(\theta)D\hat{\theta}_n} = 1.$$

граница Рао–Крамера — это наилучшая возможная (наименьшая) дисперсия среди всех несмещённых оценок при данных условиях. Если оценка достигает этой границы, она «выжала из данных максимум точности», т.е. эффективна.

Пример 1

Покажите, что оценка максимального правдоподобия параметра λ, полученная по выборке из пуассоновского закона, является эффективной.

Решение. Поскольку $P_\lambda(\xi_k = m) = \frac{\lambda^m}{m!}e^{-\lambda}$, то функция правдоподобия имеет вид:

$$\boldsymbol{L}(X_1, X_2, \ldots X_n, \lambda) = \prod_{k=1}^{n} P_\lambda(\xi_k = X_k) = \prod_{k=1}^{n} \frac{\lambda^{X_k}}{X_k!}e^{-\lambda}$$

$$\ln \boldsymbol{L}(X_1, X_2, \ldots X_n, \lambda) = \sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \ln\left(\frac{\lambda^{X_k}}{X_k!}e^{-\lambda}\right) = \sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \left(X_k \ln\lambda - \ln X_k! - \lambda\right)$$

$$\frac{\partial \ln \boldsymbol{L}(X_1, X_2, \ldots X_n, \lambda)}{\partial\lambda} = \sum_{k=1}^{n}\left(\frac{X_k}{\lambda} - 1\right) = \frac{1}{\lambda}\left(\sum_{k=1}^{n} X_k - n\right) \to \hat{\lambda}_n = \frac{1}{n}\sum_{k=1}^{n} X_k,$$

то есть оценка максимального правдоподобия совпадает с оценкой по методу моментов.

$$D\hat{\lambda}_n = \frac{1}{n^2}\sum_{k=1}^{n} DX_k = \frac{\lambda}{n}$$

Вычислим информацию Фишера:

$$\ln P_\lambda\left(\xi = X_k\right) = \ln \frac{\lambda^{X_k}}{X_k!} e^{-\lambda} = X_k \ln \lambda - \ln X_k! - \lambda, \quad \frac{\partial \ln P_\lambda\left(\xi = X_k\right)}{\partial \lambda} = \frac{X_k}{\lambda} - 1,$$

$$I_1(\lambda) = M\left(\frac{\partial \ln P_\lambda\left(\xi = X_k\right)}{\partial \lambda}\right)^2 = M\left(\frac{X_k}{\lambda} - 1\right)^2 = \frac{1}{\lambda^2} M\left(X_k - \lambda\right)^2 = \frac{1}{\lambda^2} DX_k = \frac{1}{\lambda}.$$

Следовательно, $e\left(\hat{\lambda}_n\right) = \dfrac{1}{D\hat{\lambda}_n \cdot n \cdot I_1(\lambda)} = \dfrac{1}{\frac{\lambda}{n} \cdot n \cdot \frac{1}{\lambda}} = 1.$

Пример 2

Постройте ОМП для параметра α по выборке $X_1, X_2, \ldots. X_n$ из распределения с плотностью

$$p(x,\theta) = \begin{cases} \dfrac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Параметр λ считать известным. Покажите, что полученная оценка $\hat{\alpha}$ смещенная. Постройте несмещенную оценку $\tilde{\alpha}$ параметра α. Вычислите информацию Фишера и покажите, что $\tilde{\alpha}$ асимптотически эффективна.

Составим функцию правдоподобия

$$L\left(X_1, X_2, \ldots. X_n, \alpha\right) = \prod_{k=1}^{n} \frac{\alpha^\lambda}{\Gamma(\lambda)} X_k^{\lambda-1} e^{-\alpha X_k}$$

Прологарифмируем функцию правдоподобия, чтобы было удобнее искать максимум.

$$\ln L = \sum_{k=1}^{n} \ln\left(\frac{\alpha^\lambda}{\Gamma(\lambda)} X_k^{\alpha-1} e^{-\alpha X_k}\right) = \sum_{k=1}^{n} \left(\lambda \ln \alpha - \ln \Gamma(\lambda) + (\lambda - 1) \ln X_k - \alpha X_k\right)$$

Функция правдоподобия зависит от одной переменной - α (по условию λ считается известным). Найдём максимум функции, для этого возьмём производную по α

$$\frac{\partial \ln L\left(X_1, X_2, \ldots. X_n, \lambda\right)}{\partial \alpha} = \sum_{k=1}^{n} \frac{\lambda}{\alpha} - X_k$$

Приравняем производную к нулю и найдем максимум

$$\sum_{k=1}^{n}\frac{\lambda}{\alpha}-X_k=0\Rightarrow\frac{n\lambda}{\alpha}=\sum_{k=1}^{n}X_k\Rightarrow\hat{\alpha}=\frac{\lambda}{\overline{X}}$$

Покажем, что полученная оценка является смещённой, для этого найдём её математическое ожидание.

$$M\hat{\alpha}=M\left(\frac{\lambda}{\overline{X}}\right)=\lambda nM\left(\frac{1}{X_1+X_2+\ldots X_n}\right)$$

Для того, чтобы найти математическое ожидание, нужно знать как распределена статистика $X_1+X_2+\ldots X_n$. В этой задаче было доказано, что сумма независимых случайных величин, распределенных по закону $\gamma_{\alpha,\lambda}$ и $\gamma_{\alpha,\nu}$ соответственно, распределена по закону $\gamma_{\alpha,\,\lambda+\nu}$.

Так как каждый элемент выборки распределён по закону $\gamma_{\alpha,\lambda}$, то получаем

$$\underbrace{\gamma_{\alpha,\lambda}+\gamma_{\alpha,\lambda}+\ldots\gamma_{\alpha,\lambda}}_{n}=\gamma_{\alpha,n\lambda}$$

Плотность такого распределения имеет вид

$$p_{\gamma_{\alpha,n\lambda}}(x)=\begin{cases}\dfrac{\alpha^{n\lambda}}{\Gamma(n\lambda)}x^{n\lambda-1}e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases}$$

Теперь найдём математическое ожидание полученной оценки

$$M\hat{\alpha}=\lambda nM\left(\frac{1}{X_1+X_2+\ldots X_n}\right)=\lambda n\int_0^\infty\frac{1}{x}\frac{\alpha^{n\lambda}}{\Gamma(n\lambda)}x^{n\lambda-1}e^{-\alpha x}dx=\frac{\lambda n\alpha^2}{\Gamma(n\lambda)}\int_0^\infty\alpha^{n\lambda-2}x^{n\lambda-2}e^{-\alpha x}dx=$$

$$=\left\{\begin{aligned}\alpha x&=t\\ dx&=\frac{1}{\alpha}dt\end{aligned}\right\}=\frac{\lambda n\alpha^2}{\Gamma(n\lambda)}\int_0^\infty t^{n\lambda-2}e^{-t}\frac{1}{\alpha}dt=\frac{\lambda n\alpha}{\Gamma(n\lambda)}\Gamma(n\lambda-1)=\frac{\lambda n\alpha}{\Gamma(n\lambda-1+1)}\Gamma(n\lambda-1)=$$

$$=\frac{\lambda n\alpha}{(n\lambda-1)\Gamma(n\lambda-1)}\Gamma(n\lambda-1)=\frac{\lambda n}{(n\lambda-1)}\alpha=\frac{n}{\left(n-\dfrac{1}{\lambda}\right)}\alpha$$

Таким образом полученная оценка является смещённой, но будет асимптотически несмещённой (при $n\to\infty$). Построим несмещённую оценку параметра, используя полученную оценку по методу максимального правдоподобия. Для этого умножим $\hat{\alpha}=\frac{\lambda}{\overline{X}}$ на коэффициент

$$\frac{(n\lambda-1)}{\lambda n}$$

$$\tilde{\alpha}=\frac{(n\lambda-1)}{\lambda n}\hat{\alpha}=\frac{(n\lambda-1)}{\lambda n}\frac{\lambda}{\bar{X}}=\frac{(n\lambda-1)}{n\bar{X}}$$

Математическое ожидание такое оценки будет равно α

Докажем эффективность этой оценки, для этого вычислим дисперсию оценки

$$D\tilde{\alpha}=D\left(\frac{(n\lambda-1)}{n\bar{X}}\right)=(n\lambda-1)^2 D\left(\frac{1}{\sum_{k=1}^{n}X_k}\right)$$

Как было показано ранее, статистика $\frac{1}{\sum_{k=1}^{n}X_k}$, распределена по закону $\gamma_{\alpha,n\lambda}$

, математическое ожидание $M\left(\frac{1}{\sum_{k=1}^{n}X_k}\right)=\frac{\alpha}{(n\lambda-1)}$ было вычислено ранее,

найдём $M\left(\frac{1}{\sum_{k=1}^{n}X_k}\right)^2$

$$M\left(\frac{1}{\sum_{k=1}^{n}X_k}\right)^2=\int_0^\infty\frac{1}{x^2}\frac{\alpha^{n\lambda}}{\Gamma(n\lambda)}x^{n\lambda-1}e^{-\alpha x}dx=\frac{\alpha^3}{\Gamma(n\lambda)}\int_0^\infty\alpha^{n\lambda-3}x^{n\lambda-3}e^{-\alpha x}dx=$$

$$=\left\{\begin{matrix}\alpha x=t\\ dx=\frac{1}{\alpha}dt\end{matrix}\right\}=\frac{\alpha^3}{\Gamma(n\lambda)}\int_0^\infty t^{n\lambda-3}e^{-t}\frac{1}{\alpha}dt=\frac{\alpha^2}{\Gamma(n\lambda)}\Gamma(n\lambda-2)=\frac{\alpha^2}{\Gamma(n\lambda-2+2)}\Gamma(n\lambda-2)=$$

$$=\frac{\alpha^2}{(n\lambda-2+1)(n\lambda-2)\Gamma(n\lambda-2)}\Gamma(n\lambda-2)=\frac{\alpha^2}{(n\lambda-1)(n\lambda-2)}$$

Теперь найдём дисперсию $D\left(\frac{1}{\sum_{k=1}^{n} X_k}\right)$

$$D\left(\frac{1}{\sum_{k=1}^{n} X_k}\right) = \frac{\alpha^2}{(n\lambda-1)(n\lambda-2)} - \left(\frac{\alpha}{(n\lambda-1)}\right)^2 = \frac{\alpha^2}{(n\lambda-1)^2(n\lambda-2)}$$

Тогда окончательно дисперсия оценки $\tilde{\alpha}$

$$D\tilde{\alpha} = (n\lambda-1)^2 D\left(\frac{1}{\sum_{k=1}^{n} X_k}\right) = \frac{\alpha^2}{(n\lambda-2)}$$

При $n \to \infty$, дисперсия оценки будет стремиться у нулю, следовательно полученная оценка является состоятельной

Проверим эффективность полученной оценки. Вычислим информацию Фишера, для начала возьмём логарифм от плотности распределения

$$\ln p(x,\alpha) = \ln\left(\frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x}\right) = \lambda \ln\alpha - \ln\Gamma(\lambda) - (\lambda-1)\ln x - \alpha x,$$

Продифференцируем полученное выражение по α

$$\frac{\partial \ln p(x,\alpha)}{\partial \alpha} = \frac{\lambda}{\alpha} - x$$

Найдём информацию Фишера

$$I_1(\lambda) = M\left(\frac{\partial \ln p(x,\lambda)}{\partial \lambda}\right)^2 = M\left(x - \frac{\lambda}{\alpha}\right)^2$$

Ранее здесь были вычислены характеристики гама распределения. Заметим, что математическое ожидание гамма распределения $M\xi = \frac{\lambda}{\alpha}$. Математическое ожидание квадрата отклонения случайной величины от её математического ожидания равняется дисперсии, то есть

$$M\left(x-\frac{\lambda}{\alpha}\right)^2=D\xi=\frac{\lambda}{\alpha^2}$$

$$I_1(\alpha)=\frac{\lambda}{\alpha^2}$$

Теперь найдём эффективность

$$e(\tilde{\alpha})=\frac{1}{D\tilde{\alpha}_n\cdot n\cdot I_1(\alpha)}=\frac{1}{\frac{\alpha^2}{(n\lambda-2)}\cdot n\cdot\frac{\lambda}{\alpha^2}}=\frac{(n\lambda-2)}{n\lambda}$$

Так как $n\to\infty$ эффективность стремится к 1, то полученная оценка является асимптотически эффективной.

Пример 3

Постройте ОМП для параметра λ по выборке $X_1,X_2,\ldots.X_n$ из логнормального распределения с плотностью

$$p(x,\theta)=\begin{cases}\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}} & x>0\\ 0 & x\le 0\end{cases}.$$

Параметр σ считать известным. Вычислите информацию Фишера и покажите, что полученная оценка эффективна.

Составим функцию правдоподобия

$$L(X_1,X_2,\ldots.X_n,\lambda)=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi}X_k\sigma}e^{-\frac{(\ln X_k-\lambda)^2}{2\sigma^2}}$$

Прологарифмируем функцию правдоподобия, чтобы было удобнее искать максимум.

$$\ln L(X_1,X_2,\ldots.X_n,\lambda)=\sum_{k=1}^{n}\ln\left(\frac{1}{\sqrt{2\pi}X_k\sigma}e^{-\frac{(\ln X_k-\lambda)^2}{2\sigma^2}}\right)=\sum_{k=1}^{n}\left(-\frac{1}{2}\ln 2\pi-\ln X_k\sigma-\frac{(\ln X_k-\lambda)^2}{2\sigma^2}\right)$$

Функция правдоподобия зависит от одной переменной - λ (по условию σ

считается известным). Найдём максимум функции, для этого возьмём производную по λ.

$$\frac{\partial \ln L\left(X_1, X_2, \ldots .X_n, \lambda\right)}{\partial \lambda}=\sum_{k=1}^{n}-\frac{2\left(\ln X_k-\lambda\right)}{2\sigma^2}$$

Приравняем производную к нулю и найдем максимум

$$\sum_{k=1}^{n}-\frac{2\left(\ln X_k-\lambda\right)}{2\sigma^2}=0 \Rightarrow \sum_{k=1}^{n}\left(\ln X_k-\lambda\right)=0 \Rightarrow n\lambda=\sum_{k=1}^{n}\left(\ln X_k\right) \Rightarrow$$

$$\Rightarrow \hat{\lambda}=\frac{1}{n}\sum_{k=1}^{n}\left(\ln X_k\right)$$

Докажем эффективность оценки, для этого вычислим дисперсию оценки

$$D\hat{\lambda}=D\left(\frac{1}{n}\sum_{k=1}^{n}\left(\ln X_k\right)\right)=\frac{1}{n^2}\sum_{k=1}^{n}D\left(\ln X_k\right)$$

Отдельно вычислим $D\left(\ln X_i\right)$

$$M\left(\ln X_i\right)=\int_0^{\infty}\ln x\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}dx=\int_0^{\infty}\ln x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}d\ln x=\lambda$$

При вычислении этого интеграла внесли $\frac{1}{x}$ под знак дифференциала и получили интеграл равный интегралу при вычислении математического ожидания нормального закона.

$$M\left(\ln^2 X_i\right)=\int_0^{\infty}\ln^2(x)\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}dx=\sigma^2+\lambda^2$$

При вычислении этого интеграла также внесли $\frac{1}{x}$ под знак дифференциала и получили интеграл равный интегралу при вычислении математического ожидания от квадрата нормальной величины. Найдём дисперсию

$$D\left(\ln X_i\right)=\sigma^2+\lambda^2-\lambda^2=\sigma^2$$

Теперь можем найти дисперсию оценки $\hat{\lambda}$

$$D\hat{\lambda}_n=\frac{1}{n^2}\sum_{k=1}^{n}D\left(\ln X_k\right)=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

Вычислим информацию Фишера, возьмём логарифм от плотности распределения.

$$\ln p(x,\lambda)=\ln\left(\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}\right)=-\frac{1}{2}\ln 2\pi-\ln x-\ln\sigma-\frac{(\ln x-\lambda)^2}{2\sigma^2},$$

Продифференцируем полученное выражение по λ

$$\frac{\partial\ln p(x,\lambda)}{\partial\lambda}=\frac{2(\ln x-\lambda)}{2\sigma^2}$$

Найдём информацию Фишера

$$I_1(\lambda)=M\left(\frac{\partial\ln p(x,\lambda)}{\partial\lambda}\right)^2=M\left(\frac{2(\ln x-\lambda)}{2\sigma^2}\right)^2=$$

$$=\int_0^\infty\frac{(\ln x-\lambda)^2}{\sigma^4}\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}dx=$$

$$\frac{1}{\sigma^4}\int_0^\infty\ln^2 x\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}dx-\frac{2\lambda}{\sigma^4}\int_0^\infty\ln x\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}dx+$$

$$+\frac{\lambda^2}{\sigma^4}\int_0^\infty\frac{1}{\sqrt{2\pi}x\sigma}e^{-\frac{(\ln x-\lambda)^2}{2\sigma^2}}dx$$

Здесь первый и второй интеграл уже был вычислен ранее, когда мы находили $D(\ln X_i)$ а третий, после вынесения констант, это интеграл от плотности логнормального распределения, то есть равен 1, таким образом

$$I_1(\lambda)=\frac{\sigma^2+\lambda^2}{\sigma^4}-\frac{2\lambda^2}{\sigma^4}+\frac{\lambda^2}{\sigma^4}=\frac{1}{\sigma^2}$$

Теперь найдём эффективность

$$e(\hat{\lambda}_n)=\frac{1}{D\hat{\lambda}_n\cdot n\cdot I_1(\lambda)}=\frac{1}{\frac{\sigma^2}{n}\cdot n\cdot\frac{1}{\sigma^2}}=1.$$

Следовательно, полученная оценка является эффективной.

Достаточные статистики. Критерий факторизации. Совместный закон распределения выборочного среднего и выборочной дисперсии

Достаточные статистики

Пусть $X_1, X_2, \ldots. X_n$ - выборка из распределения, зависящего от параметра $\bar{\theta}$ (реализация случайного вектора $\bar{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$ с независимыми компонентами)

Определение. Вектор-функция $\bar{t} = \bar{t}(X_1, X_2, \ldots. X_n)$ называется достаточной статистикой для оценки параметра $\bar{\theta}$, если условная функция распределения $F_{\bar{\xi}}\left(X_1, X_2, \ldots. X_n | \bar{t}\left(\bar{\xi}\right) = \bar{t}\right)$ не зависит от $\bar{\theta}$ при любых значениях $\bar{t}$.

Замечание.

Статистика $\bar{t} = \bar{t}(X_1, X_2, \ldots. X_n)$ является достаточной для оценки параметра θ, если

Дискретный случай

Условные вероятности $P\left(\xi_1 = X_1, \xi_2 = X_2, \ldots, \xi_n = X_n | \bar{t}\left(\bar{\xi}\right) = \bar{t}\right)$ не зависят от θ при любых значениях $\bar{t}$.

Непрерывный случай

Условные плотности $p_{\bar{\xi}}\left(X_1, X_2, \ldots. X_n | \bar{t}\left(\bar{\xi}\right) = \bar{t}\right)$ не зависят от θ при любых значениях $\bar{t}$.

Достаточная статистика — это такая сводка или "выжимка" всех ваших данных, которая содержит всю необходимую информацию о нужном параметре. После того как вы эту выжимку получили, сами исходные данные становятся вам не нужны для того, чтобы делать выводы об этом параметре.

Теорема (критерий факторизации).

Статистика $\bar{t} = \bar{t}(X_1, X_2, \ldots. X_n)$ является достаточной для оценки параметра $\bar{\theta}$ тогда и только тогда, когда функция правдоподобия $L(X_1, X_2, \ldots. X_n, \theta)$ имеет вид:

$$L\left(X_1, X_2, \ldots. X_n, \bar{\theta}\right) = g\left(\bar{\theta}, \bar{t}\left(X_1, X_2, \ldots, X_n\right)\right) \cdot h\left(X_1, X_2, \ldots, X_n\right).$$

Примеры.

1. Пусть $X_1, X_2, \ldots. X_n$ - выборка из нормального закона распределения $\mathcal{N}(a, \sigma)$.

Выпишем функцию правдоподобия: $L(X_1, X_2, \ldots. X_n, \lambda) = \prod\limits_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X_k - a)^2}{2\sigma^2}} =$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left(X_k-a\right)^2\right\}=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{k=1}^{n}\left(X_k\right)^2-2a\sum_{k=1}^{n}X_k+na^2\right)\right\}$$

Она зависит от выборки только через $\sum_{k=1}^{n}X_k$ и $\sum_{k=1}^{n}X_k^2$, то есть $\left(\sum_{k=1}^{n}X_k,\sum_{k=1}^{n}X_k^2\right)$- достаточная статистика.

Замечание. Вместо пары статистик $\left(\sum_{k=1}^{n}X_k,\sum_{k=1}^{n}X_k^2\right)$ можно брать $\left(\bar{X},S^2\right)$.

В самом деле,

$$L\left(X_1,X_2,\ldots.X_n,\lambda\right)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left(X_k-a\right)^2\right\}=$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left(X_k-\bar{X}+\bar{X}-a\right)^2\right\}=$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left[\left(X_k-\bar{X}\right)^2+2\left(X_k-\bar{X}\right)\left(\bar{X}-a\right)+\left(\bar{X}-a\right)^2\right]\right\}=$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2+\left(\bar{X}-a\right)\left(n\bar{X}-n\bar{X}\right)+n\left(\bar{X}-a\right)^2\right]\right\}=$$

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2+n\left(\bar{X}-a\right)^2\right]\right\}.$$

1. Пусть $X_1,X_2,\ldots.X_n$ - выборка из закона Парето с плотностью

$$p\left(x,\theta,\alpha\right)=\begin{cases}\frac{\alpha}{\theta}\left(\frac{\theta}{x}\right)^{\alpha+1} & x\geq\theta\\ 0 & x<\theta\end{cases}.$$

Функция правдоподобия

$$\boldsymbol{L}\left(X_1,X_2,\ldots.X_n,\theta,\alpha\right)=\prod_{k=1}^{n}p\left(X_k,\theta.\alpha\right)=\prod_{k=1}^{n}\frac{\alpha}{\theta}\left(\frac{\theta}{X_k}\right)^{\alpha+1} I\left(X_k\geq\theta\right)=$$

$$=\frac{\alpha^n\theta^{\alpha n}}{\left(X_1\cdot X_2\cdot\ldots\cdot X_n\right)^{\alpha+1}}\prod_{k=1}^{n}I\left(X_k\geq\theta\right)=\frac{\alpha^n\theta^{\alpha n}}{\left(X_1\cdot X_2\cdot\ldots\cdot X_n\right)^{\alpha+1}} I\left(X_{(1)}\geq\theta\right)$$

Следовательно, если $\alpha>0$ известно, то $\hat{\theta}_n=X_{(1)}$ - достаточная оценка для параметра θ.

Если же оба параметра неизвестны, то достаточно знать пару $\left(\prod_{k=1}^{n}X_k,X_{(1)}\right)$

Замечание. Вместо статистики $\prod_{k=1}^{n} X_k$ можно брать $\sum_{k=1}^{n} \ln X_k$, поскольку

$$\prod_{k=1}^{n} X_k = e^{\sum_{k=1}^{n} \ln X_k}.$$

Совместный закон распределения выборочного среднего и выборочной дисперсии для нормально распределенной генеральной совокупности.

Теорема. Если $X_1, X_2, \ldots . X_n$ –выборка из нормального распределения с параметрами a и σ, то $\bar{X}$ и $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ независимы, $\bar{X} \sim N\left(a, \frac{\sigma}{\sqrt{n}}\right)$ и $\frac{S^2(n-1)}{\sigma^2}$ имеет распределение $\chi^2(n-1)$.

Пример 4

Пусть $X_1, X_2, \ldots X_n$ - выборка из показательного закона распределения, (плотности $p_{\xi_i}(x_i) = \lambda e^{-\lambda x}, x > 0$). Найдите достаточную статистику для оценки параметра λ.

Решение. Выпишем функцию правдоподобия:

$$L(X_1, X_2, \ldots X_n, \lambda) = \prod_{k=1}^{n} \lambda e^{-\lambda X_k} = \lambda^n e^{-\lambda \sum_{k=1}^{n} X_k}$$

Она зависит от выборки только через $\sum_{k=1}^{n} X_k$, то есть $\sum_{k=1}^{n} X_k$ - достаточная статистика.

Пример 5

Найдите достаточную статистику для оценки параметра λ, если выборка $X_1, X_2, \ldots X_n$ получена из распределения Пуассона.

Пусть $X_1, X_2, \ldots X_n$ - выборка из закона Пуассона с параметром λ.

Выпишем функцию правдоподобия:

$$L(X_1, X_2, \ldots X_n, \lambda) = \prod_{k=1}^{n} \frac{\lambda^{X_k}}{X_k!} e^{-\lambda} = e^{-\lambda} \lambda^{\sum_{k=1}^{n} X_k} \prod_{k=1}^{n} \frac{1}{X_k!}$$

Применяем критерий факторизации: $h(X_1, X_2, \ldots, X_n) = \prod_{k=1}^{n} \frac{1}{X_k!}$, а

$g\left(\lambda, \sum_{k=1}^{n} X_k\right) = e^{-\lambda} \lambda^{\sum_{k=1}^{n} X_k}$ - зависит от выборки только через $\sum_{k=1}^{n} X_k$, то есть $\sum_{k=1}^{n} X_k$ - достаточная статистика для оценки параметра λ.

Пример 6

Найдите достаточную статистику для оценки параметра θ по выборке $X_1, X_2, \ldots X_n$ из распределения с плотностью $p(x, \theta) = \frac{\theta^\lambda}{\Gamma(\lambda)} x^{\lambda - 1} e^{-\theta x}$, $x > 0$. (Параметр λ считается известным.)

$$L(X_1, X_2, \ldots X_n, \lambda) = \prod_{k=1}^{n} \frac{\theta^\lambda}{\Gamma(\lambda)} X_k^{\lambda-1} e^{-\theta X_k} = \left(\frac{\theta^\lambda}{\Gamma(\lambda)}\right)^n e^{-\theta \sum_{k=1}^{n} X_k} \prod_{k=1}^{n} X_k^{\lambda-1}$$

Применяем критерий факторизации: $h\left(X_1, X_2, \ldots, X_n\right) = \prod_{k=1}^{n} X_k^{\lambda-1}$, а

$g\left(\theta, \sum_{k=1}^{n} X_k\right) = \left(\frac{\theta^\lambda}{\Gamma(\lambda)}\right)^n e^{-\theta \sum_{k=1}^{n} X_k}$ - зависит от выборки только через $\sum_{k=1}^{n} X_k$, то есть $\sum_{k=1}^{n} X_k$ - достаточная статистика для оценки параметра θ.

Ответ. $\sum_{k=1}^{n} X_k$

Пример 7

Найдите достаточную статистику для оценки параметров θ и σ по выборке $X_1, X_2, \ldots. X_n$ из распределения с плотностью $p(x, \theta, \sigma) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \theta)^2}{2\sigma^2}}$, $x > 0$.

Ответ. $\left(\sum_{k=1}^{n} \ln X_k, \sum_{k=1}^{n} \ln^2 X_k\right)$

Пример 8

Найдите достаточную статистику для оценки параметра θ по выборке $X_1, X_2, \ldots. X_n$

из распределения с плотностью $p(x, \theta) = \frac{5x^4}{\sqrt{18\pi}} e^{-\frac{\left(x^5 - 2\theta\right)^2}{18}}$

Ответ. $L\left(X_1, X_2, \ldots. X_n, \theta\right) = \prod_{k=1}^{n} \frac{5X_k^4}{\sqrt{18\pi}} e^{-\frac{\left(X_k^5 - 2\theta\right)^2}{18}} = \frac{5^n}{(18\pi)^{\frac{n}{2}}} \left(\prod_{k=1}^{n} X_k^4\right) e^{-\sum_{k=1}^{n} \frac{\left(X_k^5 - 2\theta\right)^2}{18}} =$

$$= \frac{5^n}{(18\pi)^{n/2}} \left(\prod_{k=1}^{n} X_k^4\right) e^{-\frac{1}{18} \sum_{k=1}^{n} X_k^{10}} \cdot e^{\frac{2\theta}{9} \sum_{k=1}^{n} X_k^5} \cdot e^{-\frac{2\theta^2 n}{9}}$$

Пример 9

Найдите достаточную статистику для оценки параметра p по выборке $X_1, X_2, \ldots. X_n$ из биномиального закона распределения, если число испытаний в каждой серии равно k.

Составим функцию максимального правдоподобия

$$L\left(X_1, X_2, \ldots. X_k, \theta\right) = \prod_{k=1}^{n} p\left(X_k, \theta\right) = \prod_{k=1}^{n} C_n^{X_k} \theta^{X_k} \left(1-\theta\right)^{k-X_k} = \theta^{\sum_{k=1}^{n} X_k} \left(1-\theta\right)^{kn-\sum_{k=1}^{n} X_k} \prod_{k=1}^{n} C_n^{X_k}$$

Применяем критерий факторизации: $h\left(X_1, X_2, \ldots, X_n\right) = \prod_{k=1}^{n} C_n^{X_k}$, а

$g\left(\theta, \sum_{k=1}^{n} X_k\right) = \theta^{\sum_{k=1}^{n} X_k} \left(1-\theta\right)^{kn-\sum_{k=1}^{n} X_k}$ - зависит от выборки только через $\sum_{k=1}^{n} X_k$, то есть $\sum_{k=1}^{n} X_k$ - достаточная статистика для оценки параметра θ.

Ответ. $\sum_{k=1}^{n} X_k$