Федеральное государственное автономное образовательное учреждение высшего профессионального образования

КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ

ИНСТИТУТ МАТЕМАТИКИ И МЕХАНИКИ

Специализация: Математический анализ

Специальность: 010101 – Функциональный анализ

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(дипломная работа)

Исследование последовательных статистических критериев.

Работа завершена:

«__»__________2015г ______________ Кузиков К.В

Работа проверена:

Зав. кафедрой математического анализа,
профессор, д.ф.-м.н. ______________ Насыров С.Р.

Научный руководитель,
к.ф.-м.н. ______________ Новиков П.А.

г. Казань, 2015

Содержание

§ 1. Введение

В данной дипломной работе производится построение и проверка последовательных локально наиболее мощных инвариантных критериев с распределением выводящимися из обобщенного гамма распределения.

При построении и проверке критериев используются методы последовательного анализа. Последовательный анализ исследуется относительно недавно.Дж.Эйбрехом в 1969,Р. Берком в 1975,а так же Н.Шмитцем в 1993 рассматривались случаи наблюдений последовательных независимых одинаково распределенных случайных величин; Ан.Н.Новиковым в 2009 использовался метод характеризации структуры последовательного критерия проверки простой гипотезы при простой альтернативе,идея которого заключается в рассмотрении задачи построения отимального критерия, как задачи оптимизации;П.А.Новиковым в 2010 рассматривались локально наиболее мощные последовательные критерии для случайных процессов с дискретным временем в случае независимых наблюдений и методы построения локально наиболее мощного критерия в случае многомерного параметра.Основываясь на вышесказанном,проверка распределений полученных в §3 является актуальной и новой.

Целью данной работы является построение последовательных локльно наиболее мощных инвариантных критериев и сравнение их с локально наиболее мощными критериями с фиксированным объемом выборки.

§ 2. О локально наиболее мощном последовательном критерии[1]

Пусть $X_1, X_2, \ldots, X_n, \ldots$ — случайный процесс с дискретным временем, распределение которого, P_θ, зависит от неизвестного параметра $\theta \in \Theta$, где Θ — открытое подмножество в R. Рассматривается задача проверки простой гипотезы $H_0 : \theta = \theta_0$ при сложной альтернативе $H_1 : \theta > \theta_0$, где $\theta_0 \in \Theta$ некоторое фиксированное значение параметра.

Пусть (ψ, ϕ) — последовательный критерий проверки с (рандомизированным) правилом остановки ψ и (рандомизированным) решающим правилом ϕ, где $\psi = (\psi_1, \psi_2, \ldots, \psi_n, \ldots), \phi = (\phi_1, \phi_2, \ldots, \phi_n, \ldots)$.На каждом этапе $n = 1, 2, \ldots$ значение $\psi_n(x_1, \ldots, x_n)$ понимается, как вероятность *остановиться и перейти к принятию решения* при условии, что эксперимент дошел до этапа n и что наблюдения процесса, полученные до этапа включительно, были $(x_1, x_2, \ldots, x_n)$. Правила $\psi_1, \psi_2, \ldots$ применяются последовательно, пока эксперимент на некотором этапе не остановится. При остановке на этапе $n \geqslant 1$ для принятия решения используется решающее правило ϕ_n. Значение $\phi_n(x_1, \ldots, x_n)$ понимается как условная вероятность *отклонить* нулевую гипотезу H_0 при наблюниях $(x_1, \ldots, x_n)$.

В соответствии с вышеописанной процедурой правило остановки ψ порождает случайную величину τ_ψ (*момент остановки*), имеющую распределение

$$P(\tau_\psi = n) = E(1 - \psi_1)(1 - \psi_2) \ldots (1 - \psi_{n-1})\psi_n$$

для всех $n = 1, 2, \ldots$.

В последнем выражении ψ_n понимается как функция случайных величин,

$$\psi_n = \psi_n(X_1, X_2, \ldots, X_n),$$

в то же самое время как изначально она определяется как функция

$$\psi_n = \psi_n(x_1, x_2, \ldots, x_n).$$

В дальнейшем следует руководствоваться следующим правилом, дающим интерпритацию ψ_n: если F_n — некоторая функция наблюдений и ее аргуметнты опущены, то:

- $F_n = F_n(X_1, \dots, X_n)$, если F_n стоит под знаком вероятности или математического ожидания
- $F_n = F_n(x_1, \dots, x_n)$ в противном случае.

Продолжительность последовательного эксперимента характеризуется величиной *среднего объема наблюдений*:

$$\mathcal{N}(\psi) = E\tau_\psi = \begin{cases} \sum_{n=1}^{\infty} nP(\tau_\psi = n), \text{если } P(\tau_\psi < \infty) = 1, \\ \infty, \text{в противном случае.} \end{cases}$$

Функция мощности последовательного критерия (ψ, ϕ) в точке $\theta \in \Theta$ определяется как

$$\beta(\psi, \phi) = P_\theta(\text{отклонить} H_0) = \sum_{n=1}^{\infty} E_\theta(1-\psi_1)\dots(1-\psi_{n-1})\psi_n\phi_n.$$

Здесь и далее $E_\theta(\cdot)$ означает математическое ожидание отностительно распределения P_θ процесса $X_1, X_2, \dots$.

Вероятность ошибки I рода критерия (ψ, ϕ) определяется как

$$\alpha(\psi, \phi) = \beta(\psi, \phi)$$

. Поскольку выражение типа $(1-\psi_1)\dots(1-\psi_{n-1})\psi_n$ будет встречаться часто введем для него следующее обозначение:

$$s_n^\psi = (1-\psi_1)\dots(1-\psi_{n-1})\psi_n, n = 1, 2, \dots$$

. Обозначим так же

$$ts_n^\psi = (1-\psi_1)\dots(1-\psi_{n-1}), n = 1, 2, \dots$$

$(s_1^\psi \equiv \psi_1)$ и $t_1^\psi \equiv 1$ по определению. Также введем обозначения

$$S_n^\psi = \{(x_1, \dots, x_n) : s_n^\psi(x_1, \dots, x_n) > 0\}$$

$$T_n^{\psi}=\{(x_1,\ldots,x_n):t_n^{\psi}(x_1,\ldots,x_n)>0\}$$

Предположим, что X_j имеет "функцию плотности"$f_{\theta,j}$(производная Радона-Никодима ее распределения) относительно некоторой сигма-конечной меры μ на пространстве "значений"$X_j, j=1,2,\ldots$.

Вследствии независимости наблюдений для любого $n=1,2,\ldots$ "вектор"$(X_1,\ldots,X_n)$ первых n наблюдений имеейт "плотность"

$$f_{\theta}^n(x_1,\ldots,x_n)=\prod_{j=1}^{n}f_{\theta,j}(x_j)$$

отностительно произведения мер

$$\mu^n=\underbrace{\mu\otimes\mu\otimes\ldots\otimes\mu}_{n\text{ раз}}.$$

Будем предполагать (когда это потребуется) выполнения следующих условия

УСЛОВИЕ С1'. *Для любого $j\geqslant 1$ существует такая интегрируемая отностиельно меры μ функция $\dot{f_{\theta,j}}$,что*

$$\int|f_{\theta,j}-f_{\theta_0,j}-(\theta-\theta_0)\dot{f_{\theta,j}}|d\mu=\mathrm{o}(\theta-\theta_0)$$

при $\theta\to\theta_0$.

Пусть

$$I_j(\theta_0,\theta_1)=E_{\theta_0}\ln\frac{f_{\theta_0,j}(X_j)}{f_{\theta_1,j}(X_j)}$$

различающая информация по Кульбаку-Лейблеру между распределениями X_j при $\theta=\theta_0$ и $\theta=\theta_1, j=1,2,\ldots$.

УСЛОВИЕ С2'.*Существуют такие $\delta>0$ и $0<\gamma_1<\infty$, что*

$$I_j(\theta_0,\theta)/(\theta-\theta_0)^2\leqslant\gamma_1$$

для всех $j=1,2,\ldots$ и для всех $|\theta-\theta_0|\leqslant\delta$

УСЛОВИЕ С3'*Существует $0<\gamma_2<\infty$ такое, что для всех $j=1,2,\ldots$*

$$E_{\theta_0}|\frac{\dot{f_{\theta_0,j}}(X_j)}{f_{\theta_0,j}(X_j)}|\leqslant\gamma_2$$

§ 3. О различии между распределениями схожими с обощенным гамма распределением.

Обобщенное гамма распределение с функцией плотности

$$f(x,d,\delta)=\frac{1+\delta}{\Gamma((1+d)/(1+\delta))}\cdot\frac{x^d}{\theta^{(1+d)/(1+\delta)}}\cdot e^{\frac{-x^{1+\delta}}{\theta}} \tag{1}$$

введенное Стейси и после использовавшеяся многими авторами в различных моделях для изучения "времени жизни"сложных продуктов.

В критерии с набором $x_1,\dots,x_n, n\geqslant 2$ с гипотезами содержащими параметры d и δ для обобщенного гамма распределения, для инвариантных критериев с максимальным инвариантом $T=\{x_1/x_n,\dots,x_1/x_n\}$, параметр масштаба θ оказывается излишеним.

$$f(x,d,\delta)=\frac{1+\delta}{\Gamma((1+d)/(1+\delta))}\cdot\frac{x^d}{\theta^{(1+d)/(1+\delta)}}\cdot e^{\frac{-x^{1+\delta}}{\theta}}$$

$$P(X_2/X_1<t_1',\dots,X_n/X_1<t_{n-1}')=$$

$$=\int\limits_{(X_2/X_1<t_1',\dots,X_n/X_1<t_1')}\cdots\int f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1\dots dx_n$$

Сделаем замену $x_1=x, x_2=t_1x,\dots,x_n=t_{n-1}x$

$$J=\begin{vmatrix}\partial x_1/\partial x & \partial x_1/\partial t_1 & \cdots & \partial x_1/\partial t_{n-1}\\ \partial x_2/\partial x & \partial x_2/\partial t_1 & \cdots & \partial x_2/\partial t_{n-1}\\ \vdots & \vdots & \ddots & \vdots\\ \partial x_n/\partial x & \partial x_n/\partial t_1 & \cdots & \partial x_n/\partial t_{n-1}\end{vmatrix}=\begin{vmatrix}1 & 0 & \cdots & 0\\ t_1 & x & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ t_{n-1} & 0 & \cdots & x\end{vmatrix}=x^{n-1}$$

$$\int\limits_{(t_1x<t_1',\dots,t_{n-1}x<t_{n-1}')}\cdots\int f_{X_1,\dots,X_n}(x,t_1x\dots,t_{n-1}x)|x^{n-1}|dxdt_1\dots dt_{n-1}=$$

$$\int\limits_{(t_1<t_1',\dots,t_{n-1}<t_{n-1}')}\cdots\int f_{X_1,\dots,X_n}(x,t_1x\dots,t_{n-1}x)|x^{n-1}|dxdt_1\dots dt_{n-1}$$

По теореме Фубини интеграл по области можно заменить на повторный:

$$\int\limits_{-\infty}^{t_1'}\cdots\int\limits_{-\infty}^{t_{n-1}'} f_{X_1,\dots,X_n}(x,t_1x\dots,t_{n-1}x)|x^{n-1}|dxdt_1\dots dt_{n-1}=$$

$$=\int\limits_{-\infty}^{t_1'}\cdots\int\limits_{-\infty}^{t_{n-1}'} f(x)f(t_1x)\dots f(t_{n-1}x)|x^{n-1}|dxdt_1\dots dt_{n-1}$$

Тогда плотность случайной величины $X_2/X_1,\dots,X_n/X_1$ равна:

$$f(t_1,\ldots,t_{n-1})=\int\limits_{-\infty}^{\infty} f(x)f(t_1x)\ldots f(t_{n-1}x)|x^{n-1}|dx$$

И так как $f\geq 0$:

$$f(t_1,\ldots,t_{n-1})=\int\limits_{0}^{\infty} f(x)f(t_1x)\ldots f(t_{n-1}x)|x^{n-1}|dx=$$

$$=\int\limits_{0}^{\infty}\left(\frac{1+\delta}{\Gamma((1+d)/(1+\delta))}\cdot\frac{x^d}{\theta^{(1+d)/(1+\delta)}}\cdot e^{\frac{-x^{1+\delta}}{\theta}}\right)\cdot$$

$$\cdot\left(\frac{1+\delta}{\Gamma((1+d)/(1+\delta))}\cdot\frac{(t_1x)^d}{\theta^{(1+d)/(1+\delta)}}\cdot e^{\frac{-(t_1x)^{1+\delta}}{\theta}}\right)\ldots$$

$$\ldots\left(\frac{1+\delta}{\Gamma((1+d)/(1+\delta))}\cdot\frac{(t_{n-1}x)^d}{\theta^{(1+d)/(1+\delta)}}\cdot e^{\frac{-(t_{n-1}x)^{1+\delta}}{\theta}}\right)|x^{n-1}|dx=$$

$$=\frac{(1+\delta)^n}{\Gamma^n((1+d)/(1+\delta))}\cdot\frac{\prod\limits_{i=1}^{n-1}t_i^d}{\theta^{n(1+d)/(1+\delta)}}\cdot\int\limits_{0}^{+\infty}x^{n(1+d)-1}\cdot e^{\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}{\theta}}dx=$$

$$=\ -\frac{(1+\delta)^n}{\Gamma^n((1+d)/(1+\delta))}\ \cdot\ \frac{\prod\limits_{i=1}^{n-1}t_i^d}{\theta^{n(1+d)/(1+\delta)}}\ \cdot\ \int\limits_{0}^{+\infty}-x^{n(1+d)-1}\ \cdot\ x^{-\delta}\ \cdot\ \left(-\frac{1}{1+\delta}\right)\ \cdot$$

$$\frac{\theta}{\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}\cdot e^{\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}{\theta}}\cdot d\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}{\theta}=$$

$$=\qquad\frac{(1+\delta)^{n-1}}{\Gamma^n((1+d)/(1+\delta))}\qquad\cdot\qquad\frac{\prod\limits_{i=1}^{n-1}t_i^d}{\theta^{\frac{n(1+d)}{1+\delta}-1}}\qquad\cdot\qquad\frac{1}{\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}\qquad\cdot$$

$$\int\limits_{0}^{+\infty}\left(\frac{\left(-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)\right)}{\theta}\right)^{\frac{n(1+d)-1-\delta}{1+\delta}}\cdot$$

$$\frac{\theta^{\frac{n(1+d)-1-\delta}{1+\delta}}}{\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)^{\frac{n(1+d)-1-\delta}{1+\delta}}}\cdot e^{\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}{\theta}}\cdot d\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1}t_i^{1+\delta}\right)}{\theta}=$$

$$= \frac{(1+\delta)^{n-1}}{\Gamma^n((1+d)/(1+\delta))} \cdot \frac{\prod\limits_{i=1}^{n-1} t_i^d}{\left(1+\sum\limits_{i=1}^{n-1} t_i^{1+\delta}\right)^{\frac{n(1+d)}{1+\delta}}} \cdot$$

$$\int\limits_0^{+\infty} \left(\frac{\left(-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1} t_i^{1+\delta}\right)\right)}{\theta} \right)^{\frac{n(1+d)-1-\delta}{1+\delta}} \cdot e^{\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1} t_i^{1+\delta}\right)}{\theta}} \cdot$$

$$d\frac{-x^{1+\delta}\left(1+\sum\limits_{i=1}^{n-1} t_i^{1+\delta}\right)}{\theta} =$$

$$= \frac{(1+\delta)^{n-1}\Gamma(n(1+d)/(1+\delta))}{\Gamma^n((1+d)/(1+\delta))} \cdot \frac{\prod\limits_{i=1}^{n-1} t_i^d}{\left(1+\sum\limits_{i=1}^{n-1} t_i^{1+\delta}\right)^{\frac{n(1+d)}{1+\delta}}}$$

таким образом функция плотности избавляется от параметра θ

$$p_{d,\delta}(t_1,\ldots,t_{n-1}) = \frac{(1+\delta)^{n-1}\Gamma(n(1+d)/(1+\delta))}{\Gamma^n((1+d)/(1+\delta))} \cdot \frac{\prod\limits_{i=1}^{n-1} t_i^d}{\left(1+\sum\limits_{i=1}^{n-1} t_i^{1+\delta}\right)^{\frac{n(1+d)}{1+\delta}}} \qquad (2)$$

Строя гипотезы с параметрами d и δ получить:

3.1. Обобщенное экспонециальное распределение

$d = 0$

$H_0 : \delta = 0$

$H_1 : \delta > 0$

$\frac{x_k}{x_n} = t_k$

$p_{0,\delta}(t_1,\ldots,t_{n-1})$

$$= \frac{(1+\delta)^{n-1}\Gamma(\frac{n}{1+\delta})}{\Gamma^n(\frac{1}{1+\delta})\left(1+\sum\limits_{k=1}^{n-1}(t_k)^{1+\delta}\right)^{\frac{n}{1+\delta}}}$$

$\ln p_{0,\delta}(t_1,\ldots,t_{n-1})$

$$= \ln \frac{(1+\delta)^{n-1}\Gamma(\frac{n}{1+\delta})}{\Gamma^n(\frac{1}{1+\delta})\left(1+\sum\limits_{k=1}^{n-1}(t_k)^{1+\delta}\right)^{\frac{n}{1+\delta}}}$$

$$= (n-1)\ln(1+\delta) + \ln\Gamma(\frac{n}{1+\delta}) - n\ln\Gamma(\frac{1}{1+\delta}) - \frac{n}{1+\delta}\ln\left(1+\sum_{k=1}^{n-1}(t_k)^{1+\delta}\right)$$

$$\ln p_{d,d}\left(\frac{x_1}{x_n},\ldots,\frac{x_{n-1}}{x_n}\right)$$

$$= (n-1)\ln(1+\delta) + \ln\Gamma(\frac{n}{1+\delta}) - n\ln\Gamma(\frac{1}{1+\delta}) - \frac{n}{1+\delta}\ln\left(1+\sum_{k=1}^{n-1}\left(\frac{x_k}{x_n}\right)^{1+\delta}\right)$$

$$= (n-1)\ln(1+\delta) + \ln\Gamma(\frac{n}{1+\delta}) - n\ln\Gamma(\frac{1}{1+\delta}) - \frac{n}{1+\delta}\ln\left(\sum_{k=1}^{n}\left(\frac{x_k}{x_n}\right)^{1+\delta}\right)$$

$$= (n-1)\ln(1+\delta) + \ln\Gamma(\frac{n}{1+\delta}) - n\ln\Gamma(\frac{1}{1+\delta}) - \frac{n}{1+\delta}\left(\ln\left(\sum_{k=1}^{n}(x_k)^{1+\delta}\right) - (1+\delta)\ln x_n\right)$$

$$= (n-1)\ln(1+\delta) + \ln\Gamma(\frac{n}{1+\delta}) - n\ln\Gamma(\frac{1}{1+\delta}) - \frac{n}{1+\delta}\ln\left(\sum_{k=1}^{n}(x_k)^{1+\delta}\right) + n\ln x_n$$

$$\frac{\ln p_{d,d}(\frac{x_1}{x_n},\ldots,\frac{x_{n-1}}{x_n})}{\partial\delta}$$

$$= \frac{n-1}{1+\delta} - \frac{n}{(1+\delta)^2}(\psi(\frac{n}{1+\delta}) - \psi(\frac{1}{1+\delta})) - \frac{n}{(1+\delta)^2}\ln\sum_{k=1}^{n}(x_k)^{1+\delta} + \frac{n\sum\limits_{k=1}^{n}((x_k)^{1+\delta}\ln x_k)}{(1+\delta)\sum\limits_{k=1}^{n}(x_k)^{1+\delta}}$$

$d = \delta = 0$

$$n\left(1 - \frac{1}{n} - \sum_{k=1}^{n-1}\frac{1}{k} - \ln\sum_{k=1}^{n}(x_k) + \frac{\sum\limits_{k=1}^{n}(x_k\ln x_k)}{\sum\limits_{k=1}^{n}x_k}\right)$$

3.2. Распределение Вейбулла

$\delta = d$

$H_0 : d = 0$

$H_1 : d > 0$

$$\frac{x_k}{x_n} = t_k$$

$$p_{d,d}(t_1, \ldots, t_{n-1})$$

$$= \frac{(1+d)^{n-1}\Gamma(n)\prod\limits_{k=1}^{n-1}(t_k)^d}{\Gamma^n(1)\left(1+\sum\limits_{k=1}^{n-1}(t_k)^{d+1}\right)^n}$$

$$= \frac{(1+d)^{n-1}\Gamma(n)\prod\limits_{k=1}^{n-1}(t_k)^d}{\left(1+\sum\limits_{k=1}^{n-1}(t_k)^{d+1}\right)^n}$$

$$\ln p_{d,d}(t_1, \ldots, t_{n-1})$$

$$= \ln\frac{(1+d)^{n-1}\Gamma(n)\prod\limits_{k=1}^{n-1}(t_k)^d}{\left(1+\sum\limits_{i=1}^{n-1}(t_k)^{d+1}\right)^n}$$

$$= (n-1)\ln(1+d) + \ln\Gamma(n) + d\sum_{k=1}^{n-1}\ln t_k - n\ln\left(1+\sum_{k=1}^{n-1}(t_k)^{d+1}\right)$$

$$\ln p_{d,d}\left(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}\right)$$

$$= (n-1)\ln(1+d) + \ln\Gamma(n) + d\sum_{k=1}^{n-1}\ln\frac{x_k}{x_n} - n\ln\left(1+\sum_{k=1}^{n-1}\left(\frac{x_k}{x_n}\right)^{d+1}\right)$$

$$= (n-1)\ln(1+d) + \ln\Gamma(n) + d\sum_{k=1}^{n}\ln\frac{x_k}{x_n} - n\ln\left(\frac{\sum\limits_{k=1}^{n}(x_k)^{d+1}}{(x_n)^{d+1}}\right)$$

$$= (n-1)\ln(1+d) + \ln\Gamma(n) + d\sum_{k=1}^{n}\ln\frac{x_k}{x_n} - n\ln(\sum_{k=1}^{n}(x_k)^{d+1}) + n(d+1)\ln(x_n)$$

$$\frac{\ln p_{d,d}(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n})}{\partial d}$$

$$= \frac{n-1}{1+d} + \sum_{k=1}^{n}\ln x_k - n\ln x_n - \frac{n(\ln\sum\limits_{k=1}^{n}(x_k)^{d+1})}{\partial d} + n\ln x_n$$

$$= \frac{n-1}{1+d} + \sum_{k=1}^{n}\ln x_k - \frac{n}{\sum\limits_{k=1}^{n}(x_k)^{d+1}} \cdot \sum_{k=1}^{n}\frac{(x_k)^{d+1}}{\partial d}$$

$$= \frac{n-1}{1+d} + \sum_{k=1}^{n} \ln x_k - n \frac{\sum\limits_{k=1}^{n} ((x_k)^{d+1} \ln x_k)}{\sum\limits_{k=1}^{n} (x_k)^{d+1}}$$

$d = \delta = 0$

$$n - 1 + \sum_{k=1}^{n} \ln x_k - n \frac{\sum\limits_{k=1}^{n} (x_k \ln x_k)}{\sum\limits_{k=1}^{n} x_k}$$

3.3. Гамма распределение

$\delta = 0$

$H_0 : d = 0$

$H_1 : d > 0$

$\frac{x_k}{x_n} = t_k$

$p_{d,0}(t_1, \ldots, t_{n-1})$

$$= \frac{\Gamma(n(1+d)) \prod\limits_{k=1}^{n-1} (t_k)^d}{\Gamma^n(1+d) \left(1 + \sum\limits_{k=1}^{n-1} t_k\right)^{n(1+d)}}$$

$\ln p_{d,0}(t_1, \ldots, t_{n-1})$

$$= \ln \frac{\Gamma(n(1+d)) \prod\limits_{k=1}^{n-1} (t_k)^d}{\Gamma^n(1+d) \left(1 + \sum\limits_{k=1}^{n-1} t_k\right)^{n(1+d)}}$$

$$= \ln \Gamma(n(1+d)) + d \sum_{k=1}^{n-1} \ln t_k - n \ln \Gamma(1+d) - n(1+d) \ln \left(1 + \sum_{k=1}^{n-1} t_k\right)$$

$\ln p_{d,0} \left(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}\right)$

$$= \ln \Gamma(n(1+d)) + d \sum_{k=1}^{n-1} \ln \frac{x_k}{x_n} - n \ln \Gamma(1+d) - n(1+d) \ln \left(1 + \sum_{k=1}^{n-1} \frac{x_k}{x_n}\right)$$

$$= \ln \Gamma(n(1+d)) + d \sum_{k=1}^{n} \ln \frac{x_k}{x_n} - n \ln \Gamma(1+d) - n(1+d) \ln \left(\sum_{k=1}^{n} \frac{x_k}{x_n}\right)$$

$$= \ln \Gamma(n(1+d)) + d \left(\sum_{k=1}^{n} \ln x_k - n \ln x_n\right) - n \ln \Gamma(1+d) - n(1+d)(\ln \left(\sum_{k=1}^{n} x_k\right) - \ln x_n)$$

$$\frac{\ln p_{d,0}(\frac{x_1}{x_n}, \dots, \frac{x_{n-1}}{x_n})}{\partial d} = n\psi(n(1+d)) + \sum_{k=1}^{n} \ln x_k - n\psi(1+d) - n \ln \left(\sum_{k=1}^{n} x_k \right)$$

$d = \delta = 0$

$$n \sum_{k=1}^{n-1} \frac{1}{k} + \sum_{k=1}^{n} \ln x_k + (n-1)\gamma - n \ln \left(\sum_{k=1}^{n} x_k \right)$$

§ 4. Численное исследование локально наиболее мощного критерия

Рассмотрим задачу проверки гипотезы $H_0 : d = 0$ при альтернативе $H_1 : d > 0$. Зададим вероятность ошибки первого рода 0.05. Построим локально наиболее мощный последовательный инвариантный критерии(SLMP) с удовлетворяющие ограничениям вероятности на ошибки первого рода(константы A и B подберем вручную). Сравним поведение его функции мощности(power_s) с функцией мощности(power_f) локально наиболее мощным инвариантным критерием с фиксированным объемом выборки(FSSLMP)с фиксацией на основе SLMP.

Случай гамма распределения:

```
Sequential LMP(SLMP) test and Fixed Size LMP(FSSLMP) test
"In SLMP test"
       d   power_s         n      zmin      zmax
0.000000  0.061000 38.391000 -8.634576 20.568182
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
 0.00000   0.05200  38.00000 -18.17018  12.29833
"In SLMP test"
       d   power_s         n      zmin      zmax
0.100000  0.194000 52.985000 -5.459177 20.560827
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
 0.10000   0.13500  53.00000 -18.33760  16.44305
"In SLMP test"
      d  power_s        n     zmin     zmax
0.20000  0.32800 56.60600 -5.73703 20.56986
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
 0.20000   0.25900  57.00000 -12.90341  19.36488
"In SLMP test"
       d   power_s         n      zmin      zmax
0.300000  0.427000 57.079000 -6.364689 20.571909
```

```
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
0.300000  0.440000 57.000000 -7.387653 19.332819
"In SLMP test"
       d   power_s         n      zmin      zmax
0.400000  0.499000 54.343000 -4.988263 20.559791
"In FSSLMP test whith n from SLMP"
      d  power_f        n     zmin     zmax
0.40000  0.56300 54.00000 -1.76245 21.23377
"In SLMP test"
      d  power_s        n     zmin     zmax
0.50000  0.58200 56.59800 -3.98230 20.56571
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
0.500000  0.760000 57.000000 -4.274953 21.662569
"In SLMP test"
       d   power_s         n      zmin      zmax
0.600000  0.635000 54.570000 -2.968793 20.573486
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
0.600000  0.853000 55.000000 -2.865867 20.572239
"In SLMP test"
       d   power_s         n      zmin      zmax
0.700000  0.700000 55.723000 -2.963765 20.570683
"In FSSLMP test whith n from SLMP"
        d    power_f          n       zmin       zmax
0.7000000  0.9300000 56.0000000  0.7804224 22.8576352
"In SLMP test"
       d   power_s         n      zmin      zmax
0.800000  0.718000 52.972000 -2.749227 20.575980
"In FSSLMP test whith n from SLMP"
       d   power_f         n      zmin      zmax
0.800000  0.944000 53.000000  1.982941 22.358357
"In SLMP test"
```

```
        d   power_s          n      zmin      zmax
 0.900000  0.768000 53.842000 -3.858617 20.570946
 "In FSSLMP test whith n from SLMP"
        d   power_f          n      zmin      zmax
 0.900000  0.971000 54.000000  5.322048 23.788807
 "In SLMP test"
        d   power_s          n      zmin      zmax
 1.000000  0.780000 52.255000 -2.314814 20.568911
 "In FSSLMP test whith n from SLMP"
        d   power_f          n      zmin      zmax
 1.000000  0.987000 52.000000  6.511753 24.178057
```

как можно видеть значение функции мощности для FSSLMP при альтернативе растет с ростом альтернативы быстрее чем SLMP, а это значит что вероятность ошибки второго рода для SLMP меньше чем для FSSLMP. Следовательно для последовательного критерия можно провести меньшее число экспериментов, при этом вероятность ошибки второго рода(значения функции мощности) будет меньше.
Случай распределения Вейбулла

```
"In SLMP test"
        d   power_s          n      zmin      zmax
 0.000000  0.036000 56.224000 -8.322041 55.890014
"In FSSLMP test with n from SLMP"
        d   power_f          n      zmin      zmax
  0.00000   0.04900  56.00000 -36.05450  26.40496
"In SLMP test"
         d    power_s           n       zmin       zmax
  0.100000   0.277000 100.405000  -8.144106  55.984428
"In FSSLMP test with n from SLMP"
        d   power_f          n      zmin      zmax
  0.10000   0.28700 100.00000 -25.41462  44.32281
"In SLMP test"
        d   power_s          n      zmin      zmax
 0.200000  0.465000 95.985000 -8.237856 56.009532
```

```
"In FSSLMP test with n from SLMP"
        d   power_f         n      zmin      zmax
 0.200000  0.712000 96.000000 -9.909251 58.233727
"In SLMP test"
        d   power_s         n      zmin      zmax
 0.300000  0.635000 99.558000 -5.919013 56.011147
"In FSSLMP test with n from SLMP"
         d    power_f          n       zmin       zmax
  0.300000   0.956000 100.000000   7.604612  57.433689
"In SLMP test"
        d   power_s         n      zmin      zmax
 0.400000  0.702000 91.069000 -4.796798 56.004007
"In FSSLMP test with n from SLMP"
        d   power_f         n      zmin      zmax
 0.400000  0.991000 91.000000  9.335272 61.570079
"In SLMP test"
        d   power_s         n      zmin      zmax
 0.500000  0.801000 90.055000 -5.033541 56.011924
"In FSSLMP test with n from SLMP"
       d  power_f        n     zmin     zmax
 0.50000  1.00000 90.00000 23.23068 61.28617
"In SLMP test"
       d  power_s        n     zmin     zmax
 0.60000  0.82300 84.49400 -3.72954 55.99012
"In FSSLMP test with n from SLMP"
       d  power_f        n     zmin     zmax
 0.60000  1.00000 84.00000 25.35652 63.48408
"In SLMP test"
        d   power_s         n      zmin      zmax
 0.700000  0.837000 79.137000 -3.498982 56.009425
"In FSSLMP test with n from SLMP"
       d  power_f        n     zmin     zmax
 0.70000  1.00000 79.00000 26.99502 59.76693
"In SLMP test"
```

```
       d   power_s          n      zmin      zmax
 0.800000  0.890000 79.276000 -4.570484 56.002032
"In FSSLMP test with n from SLMP"
      d  power_f        n     zmin     zmax
 0.80000  1.00000 79.00000 25.33354 59.52673
"In SLMP test"
       d   power_s          n      zmin      zmax
 0.900000  0.915000 77.244000 -3.078109 55.994094
"In FSSLMP test with n from SLMP"
      d  power_f        n     zmin     zmax
 0.90000  1.00000 77.00000 34.70709 62.02961
"In SLMP test"
       d   power_s          n      zmin      zmax
 1.000000  0.935000 75.811000 -2.366296 55.990925
"In FSSLMP test with n from SLMP"
      d  power_f        n     zmin     zmax
 1.00000  1.00000 76.00000 37.60439 62.13417
```

как можно видеть значение функции мощности для FSSLMP при альтернативе растет с ростом альтернативы быстрее чем SLMP, а это значит что вероятность ошибки второго рода для SLMP меньше чем для FSSLMP. Следовательно для последовательного критерия можно провести меньшее число экспериментов, при этом вероятность ошибки второго рода(значения функции мощности) будет меньше.

§ 5. Заключение

В работе изучены и построены примеры локально наиболее мощных последовательных инвариантных критериев. В ходе построения был освоен материал, представленный в соответствующей литературе. Выведены распределениия : обобщенное эспоненциальное, гамма, Вейбулла из обобщенного гамма распределения. Построены и сравнены примеры локально наиболее мощных инвариантных последовательных критериев и локально наиболее мощных инвариантных критериев с фиксированным объемом выборки.На основе построения сделан вывод: для последовательного критерия можно провести меньшее число экспериментов, при этом вероятность ошибки второго рода(значения функции мощности) будет меньше.

Список литературы

[1] Новиков П.А. *Локально наиболее мощные критерии проверки гипотез о параметрах случайных процессов с дискретным временем.* / П.А.Новиков, Казань, 2010, 104с.

[2] Зарядов И.С. *Введение в статистический пакет R: типы переменных, структуры данных, чтение изапись информации, графика* / И.С.Зарядов, Москва, 2010, 207с.

[3] Новиков Ан.А. *Элементарное введение в последовательный статистический анализ* / Ан.А.Новиков, Мехико, 2007, 25с.

[4] Леман Э. *Проверка статистических гипотез* / Э.Леман, Москва, 1979, 404с.

[5] Volodin I.N. *On Discriminating Between Types Connected with the Generalized Gamma Distribution* / I.N.Volodin Selected Transl. in Math. Statist. and Probability, 1981, 15c.

[6] Kumar V.,Ligges U. *Package "reliaR"* / V.Kumar, U.Ligges, 2015, 153.

[7] Боровков А.А. *Математическая статиистика* / А.А.Боровков, Новосибирск:Наука, издательство института математики, 1997, 772с.

[8] Де-Гроот М. *Оптимальные статистические решения* / М.Де-Гроот, М.:Мир, 1974

[9] Новиков.П.А. Локально наиболее мощные последовательные критерии для марковсских процессов с дискретным временем. *Теория вероятностей и ее применения* / П.А.Новиков, 2010, 55(2), с.369-372.

[10] Abraham J.K. *The local power of sequentional tests subject to an expected sample size restriction* / J.K.Abraham, Unpublished Stanford technical report, 1969.

[11] Berk R.H. Locally most powerful sequentional tests. *Annals of Statistics.* / R.H.Berk. 3., 1975, P.373-381.

[12] Ghosh M., Mukhopadhyay N., Sen P.K., *Sequentional Estimation* /M.Ghosh, N.Mukhopadhyay,P.K.Sen, New-York:Wiley, 1997.

[13] Novikov A. Optimal sequential tests for two simple hypotheses. *Sequential analysis.* /A.Novikov ,2009.,28(2)., P. 188-217.

[14] Novikov A., Novikov P. Locally most powerful sequential tests of a simple hypothesis vs. one-sided alternatives. *Journal of Statistical Planning and Inference.* / A.Novikov, P.Novikov, 2010. , 140(3). , P. 750-765.

[15] Wald A., Wolfowitz J. Optimum character of the sequential probability ratio test. *Annals of Mathematical Statistics.* / A.Wald, J.Wolfowitz, 19. , 1948. , P. 326-339.

[16] Stasy E.W. *A generalization of gamma-distribution* / E.W.Stasy, Ann.Math.Statist.,28,1962,P.1187-1192.

[17] Harold W. Hager, Lee J. Bain *Inferential procedures for gamma-distribution* / W. Hager. Harold , J. Bain.Lee, J.Amer.Statist.Assoc. 332(1970), P.1601-1609.

[18] Володин И.Н., Степин А.Г., Фишбейн М.А. *О мощности одного критерия для различения двух близких типов Вейбулла при "мешающем"масштабном параметре* / И.Н. Володин ,А.Г. Степин ,М.А. Фишбейн, Ученые записки Казанского ун-та, 127, 3(1967), с.25-40.

[19] Володин И.Н. *Проверка статистических гипотез о типе распределения по малым выборкам.* / И.Н. Володин Ученые записки Казанского ун-та, 125, 6(1966), с.3-23

Приложение 1
Листинг когда R для проверки гамма распределения

```
defmaxiters <- 1000

fix.mc.power <- function(n, d, C, maxiters=defmaxiters)
{
zmin <- +Inf
zmax <- -Inf
rejected <- 0.
accepted <- 0.
for(ri in 1:maxiters)
{
x <- rgamma(n, shape=1+d, scale=4.)
z <- n*digamma(n) - n*digamma(1) + sum(log(x))
- n*log(sum(x))
zmin <- min(zmin,z)
zmax <- max(zmax,z)
if(z > C)
{
rejected <- rejected + 1
}
else
{
accepted <- accepted + 1
}
}
r <- rejected/maxiters
a <- accepted/maxiters
avgn <- n
c("d"=d,"power_f"=r,"n"=avgn,"zmin"=zmin,"zmax"=zmax)
}

fix.mc.getC <- function(alpha,n,d0,interval = c(-1000,1000))
{
obj <- function(C) fix.mc.power(n,d=d0,C)[["power_f"]]-alpha
ur <- uniroot(obj,interval)
ur$root
}

seq.mc.power <- function(d, d0, A, B, maxiters=defmaxiters)
{
rejected <- 0
```

```
accepted <- 0
sumn <- 0.
zmin <- +Inf
zmax <- -Inf
for(ri in 1:maxiters)
{
x <- c()
repeat
{
x[length(x) + 1] <- rgamma(1, shape=1+d, scale=4.)
n <- length(x)
z <- n*digamma(n*(1+d0)) - n*digamma(1.+d0)
+ sum(log(x)) n*log(sum(x))
zmin <- min(zmin,z)
zmax <- max(zmax,z)
if(z > B) #reject H0
{
rejected <- rejected + 1
break
}
if(z < A)
{
accepted <- accepted + 1
break
}
}
sumn <- sumn + length(x)
}
acc <- accepted/maxiters
rej <- rejected/maxiters
avgn <- sumn/maxiters
c("d"=d,"power_s"=rej,"n"=avgn,"zmin"=zmin,"zmax"=zmax)

}
print.noquote("Sequential LMP(SLMP) test and Fixed Size LMP(FSSLMP) te
A <- -0.1
B <- 20
for(k in lambda0)
{
a <- seq.mc.power(d=d0+k,d0=d0,A,B)
print(c("In SLMP test"))
print(a)
C <- fix.mc.getC(alpha=0.05, round(a[["n"]],0) , d0)
```

```
print(c("In FSSLMP test whith n from SLMP"))
b <- fix.mc.power(round(a[["n"]],0) ,d=d0+k,C)
print(b)
}
```

листинг когда R для проверки распределения Вейбулла

```
defmaxiters <- 1000

fix.mc.power <- function(n, d, C, maxiters=defmaxiters)
{
zmin <- +Inf
zmax <- -Inf
rejected <- 0.
for(ri in 1:maxiters)
{
x <- rweibull(n, shape=1+d, scale=4.)
z <- n-1+sum(log(x)) - n*sum(x*log(x))/sum(x)
zmin <- min(zmin,z)
zmax <- max(zmax,z)
if(z > C)
rejected <- rejected + 1
}
p <- rejected/maxiters
avgn <- n
c("d"=d,"power_f"=p,"n"=avgn,"zmin"=zmin,"zmax"=zmax)
}

fix.mc.getC <- function(alpha,n,d0,interval = c(-1000,1000))
{
obj <- function(C) fix.mc.power(n,d=d0,C)[["power_f"]]-alpha
ur <- uniroot(obj,interval)
ur$root
}

seq.mc.power <- function(d, d0, A, B, maxiters=defmaxiters)
{
rejected <- 0.
sumn <- 0.
zmin <- +Inf
zmax <- -Inf
for(ri in 1:maxiters)
{
x <- c()
```

```
repeat
{
x[length(x) + 1] <- rweibull(1, shape=1+d, scale=4.)
n <- length(x)
z <- (n-1)/(1 + d0) + sum(log(x))
- n * sum((x^(d0 + 1))* log(x))/(sum(x^(d0 + 1)))
zmin <- min(zmin,z)
zmax <- max(zmax,z)
if(z > B) #reject H0
{
rejected <- rejected + 1
break
}
if(z < A)
{
break
}
}
sumn <- sumn + length(x)
}
rej <- rejected/maxiters
avgn <- sumn/maxiters
c("d"=d,"power_s"=rej,"n"=avgn,"zmin"=zmin,"zmax"=zmax)
}
d0 <- 0.
print.noquote("Sequential LMP(SLMP) test and Fixed Size LMP(FSSLMP) te
A <- -0.4
B <- 55
for(k in lambda0)
{
a <- seq.mc.power(d=d0+k,d0=d0,A,B)
print(c("In SLMP test"))
print(a)
C <- fix.mc.getC(alpha=0.05, round(a[["n"]],0) , d0)
print(c("In FSSLMP test whith n from SLMP"))
b <- fix.mc.power(round(a[["n"]],0) ,d=d0+k,C)
print(b)
}
```