$W=\sum_i \xi_i \qquad E\xi_i=0 \qquad \sum E\xi_i^2=1$

$P(a\le W\le b)\le b-a+2\gamma$, $\gamma=\sum E|\xi_i|^3$

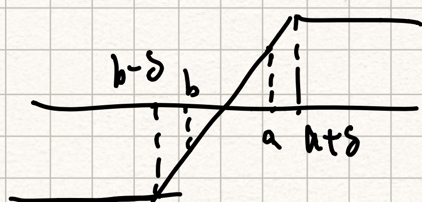
$$E(Wf(W))=\sum E\int_{-\infty}^{\infty} f'(W^{(i)}+t)\,K_i(t)\,dt$$

idea: $E(Wf(W))=\sum E(\xi_i f(W))=\sum E\xi_i\left(f(W)-f(W-\xi_i)\right)$

$$=\sum E\xi_i\int_{-\infty}^{\infty} f'(W+t)\left(1(-\xi_i<t<0)-1(0<t<-\xi_i)\right)dt$$

$$=\sum E\int_{-\infty}^{\infty} \underbrace{f'(W+t)}_{\text{hope it to be } 1(a\le W\le b)}\ \underbrace{\xi_i\left(1(\cdot\cdot)-1(\cdot\cdot)\right)}_{\hat{K}_i(t)\ \ge 0}dt$$

'hat' means random variable



构造函数

$$f(w)=\begin{cases}-\frac{b-a}{2}-\delta, & w<a-\delta\\ w-\frac{a+b}{2}, & a-\delta\le w\le b+\delta\\ \frac{b-a}{2}+\delta, & w>b+\delta\end{cases}\quad\Rightarrow\quad |f(w)|\le\frac{b-a}{2}+\delta,\quad f'(w)\ge 0$$

$$E\,Wf(W)\le\left(\frac{b-a}{2}+\delta\right)E|W|\le\frac{b-a}{2}+\delta$$ upper bound.

$$\sum E\int_{-\infty}^{\infty} f'(W+t)\,\hat{K}_i(t)\,dt\ \ge\ \sum E\,1_{(a\le W\le b)}\int_{|t|\le\delta}\hat{K}_i(t)\,dt$$

$$= \sum E \mathbb{1}(a \le W \le b) |\xi_i| \min(|\xi_i|, \delta)$$

$$= E\, \mathbb{1}(a \le W \le b) \sum \underbrace{|\xi_i| \min(|\xi_i|, \delta)}_{\eta_i}$$

$$= E\, \mathbb{1}(a \le W \le b) \sum (E\eta_i) + E\, \mathbb{1}(a \le W \le b) \sum (\eta_i - E\eta_i)$$

$$\ge P(a \le W \le b) \sum (E\eta_i) - E\left|\sum (\eta_i - E\eta_i)\right|$$

$$\le \sqrt{\sum E\eta_i^2}$$

$$\le \sqrt{\sum E\xi_i^2 \delta^2} = \delta$$

$$\frac{(b-a)}{2} + \delta \ge P(a \le W \le b) \sum E\eta_i - \delta \ge \frac{1}{2} P(a \le W \le b) - \delta$$

$$\min(x, y) \ge x - \frac{x^2}{4y}$$

$$E\eta_i \ge E\left(\xi_i^2 - \frac{|\xi_i|^3}{4\delta}\right)$$

$$\sum E\eta_i \ge \sum E\xi_i^2 - \frac{1}{4\delta} \sum E|\xi_i|^3 \qquad \text{choose } \delta = \frac{1}{2} \sum E|\xi_i|^3$$

$$= 1 - \frac{1}{2} = \frac{1}{2}$$

$$\Rightarrow P(a \le W \le b) \le b - a + 4\delta$$

$$E(\xi_i^2 \mathbb{1}(|\xi_i| \le \delta)) = E\xi_i^2 - E\xi_i^2 \mathbb{1}(|\xi_i| > \delta)$$

$$\ge E\xi_i^2 - E|\xi_i|^3 / \delta$$

① $\left|P(W\le z)-\Phi(z)\right| \le C\sum E|\xi_i|^3$

② $\left|P(W\le z)-\Phi(z)\right| \le C\sum E|\xi_i|^3/(1+|z|^3)$ non-uniform bayesian bound.

$$f(w)=\begin{cases} e^{\lambda w}(w-(a-\delta)), & a-\delta\le w<b+\delta \\ e^{\lambda w}(b-a+2\delta), & w\ge b+\delta \end{cases}$$

$f'(w)\ge e^{\lambda w}, \quad a-\delta\le w\le b+\delta$

$E e^{\lambda w}1(a\le w\le b) \ge e^{\lambda a} P(\quad)$

$T = W + \Delta$ $\qquad W = \sum \xi_i \qquad E\xi_i = 0$

$\Delta = \Delta(\xi_1, \dots, \xi_n)$

$$|P(W+\Delta \geq z) - \Phi(z)|$$

$$\leq \sup_x |P(W \leq x) - \Phi(x)| + \underbrace{2(E|\Delta|^p)^{1/(p+1)}}_{+\,4E|W\Delta|}, \quad p>0$$

– Shorack (2000). Probability for Statististians.

$$P(W+\Delta \leq z) - P(W \leq z) \leq P(\underset{\substack{\uparrow \\ \text{const.}}}{z} < W \leq \underset{\substack{\uparrow \\ \text{r.v.}}}{z-\Delta}) \leq \cdots + E|W\Delta|$$

$$+ \sum E|\xi_i(\Delta - \Delta_i)|$$

Randomized Concentration Inequality

$$P(\Delta_1 \leq W \leq \Delta_2) \leq 4\sum E|\xi_i|^3 + E|W(\Delta_2 - \Delta_1)|$$

$$+ \sum E|\xi_i(\Delta_1 - \Delta_{1,i})| + \sum E|\xi_i(\Delta_2 - \Delta_{2,i})|$$

where $\Delta_{1,i} = \Delta_{1,i}(\xi_j, i \neq j)$

$\Delta_{2,i} = \Delta_{2,i}(\xi_j, i \neq j)$

[Pf.

$$f_{a,b}(w) = \begin{cases} -\frac{b-a}{2} - \delta, & w \leq a - \delta \\ w - \frac{a+b}{2}, & a-\delta \leq w \leq b+\delta \\ \frac{b-a}{2} + \delta, & w \geq b+\delta \end{cases}$$

$\boxed{\text{Refer to Chapter 10}}$

$$EWf_{\Delta_1,\Delta_2}(W) = \sum E\xi_i\left(f_{\Delta_1\Delta_2}(W) - f_{\Delta_1\Delta_2}(W - \xi_i)\right)$$

$$+\sum E_{\xi_i}\left(f_{\Delta_1\Delta_2}(w-\xi_i)-f_{\Delta_{1i}\Delta_{2i}}(w-\xi_i)\right) \quad \#]$$

(trick) independent

Note: $\left|f_{a,b}(w)-f_{c,d}(w)\right| \le |a-c|+|b-d|$

U-Statistics

Examples: $$U_n=\frac{1}{\binom{n}{m}}\sum_{1\le i_1<i_2\cdots\le i_m\le n} h(X_{i_1},\dots,X_{i_m})$$

$$\theta=Eh(X_1,\dots,X_m)$$

$m=2$

$$h(y,x)=h(x,y) \qquad \theta=Eh(X_1,X_2)$$

Koroljuk, Borouskich (1994). Theory of U-Stat.

[Pf. W.L.O.G. Assume $\theta=0$

$$\sum_{1\le i<j\le n} h(x_i,x_j)=\sum_{j=2}^{n}\sum_{i=1}^{j-1}\underbrace{\left[h(x_i,x_j)-\underbrace{Eh(x_i,x_j|\mathcal{F}_{j-1})}_{=g(x_i)}-g(x_j)\right]}_{\text{constructed martingale difference}}$$

which has mean $g(x_j)$

$\mathcal{F}_j=\sigma(X_1,\dots,X_j)$

$g(x)=Eh(x,X_2)$

$\sigma_1^2=Var(g(X_1))$

$$+\underbrace{\sum_{j=2}^{n}\sum_{i=1}^{j-1}\left(g(x_i)+g(x_j)\right)}_{=(n-1)\sum_{i=1}^{n}g(x_i)}$$

$$U_n = \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} h(X_i, X_j)$$

$$= \frac{1}{\binom{n}{2}} \left((n-1) \sum g(X_i) + \sum_{j=2}^{n} \sum_{i=1}^{j-1} \left(h(X_i, X_j) - g(X_i) - g(X_j) \right) \right)$$

$$= \frac{2}{n} \sum g(X_i) + \frac{2}{n(n-1)} \sum_{j=2}^{n} \sum_{i=1}^{j-1} (\ldots)$$

$$\frac{\sqrt{n}}{2} \frac{U_n}{\sigma_1} = \frac{\sum g(X_i)}{\sqrt{n}\,\sigma_1} + \frac{1}{\sqrt{n}(n-1)} \sum_{j=2}^{n} \sum_{i=1}^{j-1} (\ldots)$$

$$= \sum \xi_i + \Delta$$

$$E(\Delta^2) = O\left(\frac{1}{n}\right)$$

$$E(\Delta - \Delta_i)^2 = O\left(\frac{1}{n^2}\right)$$

]