Statistics 607 **Due: Wednesday 13 April**
2005: Sheet 3

Please correct all the errors in the following outline then write a readable account of one way to get maximal inequalities for sums of independent processes using VC arguments.

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $\{\ \{f_i(t,\omega) : t \in T\} : i = 1, 2, \dots, n\}$ be independent stochastic processes. Suppose also that there are random variables F_i for which $\sup_t |f_i(t,\omega)| \le F_i(\omega)$ for all ω. Initially you should assume that T is finite, in which case each f_i could be thought of as a random vector. Let $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ be a copy of the original probability space. Let $\mathbb{Q}$ denote the uniform distribution on $\{-1,+1\}^n$. (Under $\mathbb{Q}$ the coordinate maps $\sigma_1, \dots, \sigma_n$ are independent Rademacher variables.) Define

$$
\begin{aligned}
S_t(\omega) &= \sum\nolimits_{i \le n} f_i(t,\omega) && \text{for } \omega \in \Omega \\
\widetilde{S}_t(\widetilde{\omega}) &= \sum\nolimits_{i \le n} f_i(t,\widetilde{\omega}) && \text{for } \widetilde{\omega} \in \widetilde{\Omega} \\
S_t^\circ(\omega,\widetilde{\omega},\sigma) &= \sum\nolimits_{i \le n} \sigma_i \left(f_i(t,\omega) - f_i(t,\widetilde{\omega}) \right) \\
Z_t(\omega,\sigma) &= \sum\nolimits_{i \le n} \sigma_i f_i(t,\omega)
\end{aligned}
$$

(3.1) [symmetrization of tail probabilities] Let $m_t = \text{median}(\widetilde{S}_t)$ and $m := \max_t |m_t|$. Expand the following outline to show, for each $x \ge 0$, that

$$\mathbb{P}\{\max_{t\in T} |S_t - \mathbb{P}S_t| \ge m + 2x\} \le 2\mathbb{P}\widetilde{\mathbb{P}}\{\max_{t\in T} |S_t - \widetilde{S}_t| \ge 2x\} \le 4\mathbb{P}\mathbb{Q}\{\max_{t\in T} |Z_t| \ge x\}$$

(i) Without loss of generality suppose $T = \{1, 2, \dots, N\}$ for some finite N. Let

$$
\begin{aligned}
A_t &= \{S_i - \mathbb{P}S_i \ge m_i + 2x \text{ for the first time at } i = t\} \\
B_t &= \{\widetilde{S}_t - \widetilde{\mathbb{P}}\widetilde{S}_t \le m_t\}.
\end{aligned}
$$

Show that

$$
\begin{aligned}
\mathbb{P}\{\exists t \in T : S_t - \mathbb{P}S_t \ge m_t + 2x\} &\le \sum\nolimits_t \mathbb{P}A_t(2\widetilde{\mathbb{P}}B_t) \\
&\le 2\mathbb{P}\widetilde{\mathbb{P}}\{\exists t \in T : S_t - \widetilde{S}_t \ge 2x\} \\
&\le 2\mathbb{P}\widetilde{\mathbb{P}}\mathbb{Q}\{\exists t \in T : S_t^\circ \ge 2x\} \\
&\le 4\mathbb{P}\mathbb{Q}\{\exists t \in T : Z_t \ge x\}
\end{aligned}
$$

(ii) Argue similarly for the lower tail.

(3.2) [symmetrization of Orlicz norms] Let $\Psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a convex, increasing function with $\Psi(0) < 1$. Let $\|\cdot\|_\Psi$ denote the corresponding Orlicz norm. Show that

$$\|\max_t |S_t - \mathbb{P}S_t|\,\|_\Psi \le 2\|\max_t |Z_t|\,\|_\Psi.$$

Hint: If $K = \|\max_t |S_t^\circ|\,\|_\Psi$, show that $\mathbb{P}\Psi\left(\max_t |S_t - \mathbb{P}S_t|/K\right) \le 1$.

(3.3) Let $\Psi(x) = \frac{1}{2}\exp(x^2)$. For fixed ω, identify t with the point $\mathbf{f}_t = \mathbf{f}_t(\omega) := (f_1(t,\omega), \dots, f_n(t,\omega))$ in $\mathbb{R}^n$, thereby identifying T with a subset $\mathbb{F}_\omega$ of the hypercube $\prod_i [-F_i(\omega), F_i(\omega)]$ in $\mathbb{R}^n$. Define $|F|_\omega = \left(\sum_i F_i(\omega)^2\right)^{1/2}$. Define $\delta_\omega := \sup_t |\mathbf{f}_t(\omega)|/|F|_\omega$. Note that $\delta_\omega \le 1$.

(i) For each $\alpha \ge 1$ show that there exists a finite constant C_α such that $\|X\|_\alpha \le C_\alpha \|X\|_\Psi$, for every random variable X. Hint: Consider first the case where α is an even integer.

(ii) Under $\mathbb{Q}$, with ω fixed, show that $\{Z_t : t \in T\}$ has subgaussian increments with

$$\|Z_s - Z_t\|_\Psi^2 \le |\mathbf{f}_s - \mathbf{f}_t|^2 := \sum\nolimits_i |f_i(s,\omega) - f_i(t,\omega)|^2.$$

[Do I need an extra constant on the right-hand side? If so, you woud need to adjust constants in what follows.]

(iii) Let $N_\omega(\epsilon)$ denote the smallest number of balls of (Euclidean) radius ϵ needed to cover $\mathbb{F}_\omega \cup \{0\}$. Suppose there exists a function $\lambda(\epsilon)$ for which $\int_0^1 \lambda(y)\,dy < \infty$ and

$$N(\epsilon |F|_\omega) \le \exp\left(\lambda(\epsilon)^2\right) \qquad \text{for every } \omega \text{ and } 0 < \epsilon \le 1$$

Show that there is a universal constant C_Ψ such that

$$\| \max_t |Z_t| \|_\Psi \le |F|_\omega J\left(\delta_\omega\right) \qquad \text{where } J(\epsilon) := C_\Psi \int_0^\epsilon \lambda(y)\,dy$$

(3.4) Show that

$$\mathbb{P}\{\max_t |S_t - \mathbb{P}S_t| \ge m + 2x\} \le 8\mathbb{P}\exp\left(\frac{-x^2}{|F|_\omega^2 J\left(\delta_\omega\right)^2}\right)$$
$$\le 8\mathbb{P}\{|F|_\omega > K/J(1)\} + 8\exp(-x^2/K^2) \qquad \text{for each } K > 0.$$

(3.5) For each $\alpha \ge 1$, show that there exists a finite universal constant K_α such that

$$\| \max_t |S_t - \mathbb{P}S_t| \,\|_\alpha \le K_\alpha \| \, |F|_\omega J(\delta_\omega) \, \|_\alpha \le K_\alpha J(1) \left(\mathbb{P}\left(\sum\nolimits_{i\le n} F_i^2\right)^{\alpha/2}\right)^{1/\alpha}$$

In particular, show that

$$\mathbb{P}\max_t |S_t - \mathbb{P}S_t|^2 \le (K_2 J(1))^2 \sum\nolimits_{i\le n} \mathbb{P}F_i^2$$

(3.6) What happens as T expands up to a countable dense subset of some infinite index set?

References

Pollard, D. (1989), 'Asymptotics via empirical processes (with discussion)', *Statistical Science* **4**, 341–366.

Pollard, D. (1990), *Empirical Processes: Theory and Applications*, Vol. 2 of *NSF-CBMS Regional Conference Series in Probability and Statistics*, Institute of Mathematical Statistics, Hayward, CA.