David Pollard EMAIL: pollard@stat.yale.edu
Yale University WWW: http://descartes.stat.yale.edu

Hidden convexity

1

Let $\mathbf{X} = (X_1, \ldots, X_k)$ have a multivariate normal distribution, $N(0, V)$. Put $\lambda = \text{trace} V$. Show $\mathbb{P}|\mathbf{X}|^3 \le \lambda^{3/2}\mathbb{P}|N(0,1)|^{3/2}$. ✠ Wlog $V = \text{diag}(\sigma_1^2, \ldots, \sigma_k^2)$. Convex weights $\theta_i = \sigma_i^2/\lambda$. Expectation equals $\lambda^{3/2}\mathbb{P}|\sum_i \theta_i Z_i^2|^{3/2}$ for independent standard normal Z_i. Convex function of $\boldsymbol{\theta}$ maximized at extreme point.
See Pollard (1996*b*, Chapter 13; *retrieve PROB.Coupling.ps from WWW site*) for an exposition of the application to the Yurinskii coupling.

2

Bennett's Inequality: Let $Y_1, \cdots, Y_n$ be independent random variables with (i) $\mathbb{P}Y_i = 0$ and $\mathbb{P}Y_i^2 = \sigma_i^2 < \infty$ and (ii) $Y_i \le M$ for every i, for some finite constant M. Then, for $x \ge 0$,

$$\mathbb{P}\{Y_1 + \cdots + Y_n \ge x\} \le \exp\left(-\frac{x^2}{2V}\psi\left(\frac{Mx}{V}\right)\right) \qquad \text{with } V = \sigma_1^2 + \cdots + \sigma_n^2,$$

where $\psi(x) = ((1+x)\log(1+x) - x)/(x^2/2)$ for $x \ge -1$.
Bernstein's Inequality: Replace $\psi(x)$ by its lower bound, $\psi(x) \ge (1 + \frac{x}{3})^{-1}$.
See Pollard (1996*b*, Chapter 4; *retrieve PROB.Exponential.ps from WWW site*) or Shorack & Wellner (1986, Chapter 11) for applications of the Bennett exponential inequality.

3

(Pinsker's/Csiszar's/Kemperman's inequality) Probability measures P and Q with densities p and q. Show $\int p\log(p/q) \ge 1/2\left(\int|p-q|\right)^2$. ✠ Write $p/q = 1+\delta$. Note $Q\delta = \int q(p-q)/q = 0$.

$$\begin{aligned}
\text{LHS} &= Q\left((1+\delta)\log(1+\delta) - \delta\right) && \text{because } Q\delta = 0 \\
&\ge Q\left(1/2\,\delta^2\psi(\delta)\right) && \text{definition of } \psi \\
&\ge 1/2\, Q\left(\frac{\delta^2}{1+\delta/3}\right) Q(1+\delta/3) && \text{lower bound for } \psi\text{, and } Q\delta = 0 \\
&\ge 1/2\left(Q\frac{|\delta|}{\sqrt{1+\delta/3}}\sqrt{1+\delta/3}\right)^2 && \text{Cauchy-Schwarz} \\
&= \text{RHS}
\end{aligned}$$

4

Let f be a twice differentiable convex function defined on a convex interval $J \subseteq \mathbb{R}$ that contains the origin. Suppose $f(0) = f'(0) = 0$. Use the representations

$$f(t) = t\int\{0 \le u \le 1\} f'(tu)\,du = t^2 \iint \{0 \le v \le u \le 1\} f''(tv)\,dv\,du = t^2\int_0^1 (1-v) f''(tv)\,dv$$

to establish the following facts. (i) The function $f(t)/t$ is increasing. (ii) The function $\phi(x) := 2f(t)/t^2$ is nonnegative. (iii) If f is convex then so is ϕ. (iv) If f'' is increasing then so is ϕ. (v) Invoke Jensen's inequality for the uniform distribution on $\{0 \le v \le u \le 1\}$ to show that $\phi(t) \ge f''(t/3)$.

5

Perhaps optimization estimator defined to maximize process $G_n(t) = \frac{1}{n}\sum_{i \le n} g(x_i, t)$ over t in an index set $T \subseteq \mathbb{R}$. Empirical process $\nu_n g = n^{-1/2}\sum_{i \le n}(g(x_i) - \mathbb{P}g(x_i))$.
Signal plus noise split: $G_n(t) = \mathbb{P}G_n(t) + n^{-1/2}\nu_n^x g(x,t)$.
Taylor: $g(x,t) - g(x,0) - tg'(x,0) = t\int_0^1 \left(g'(x,tu) - g'(x,0)\right)\,du$.

David Pollard: July 1995, Kingston

Stochastic Taylor: $\nu_n^x g(x,t) = \nu_n^x g(x,0) + t\nu_n^x g'(x,0) + R_n(t)$, where

$$\sup_{|t|\le\delta} \frac{|R_n(t)|}{|t|} = \sup_{|t|\le\delta} |\nu_n^x \int_0^1 \left(g'(x,tu) - g'(x,0)\right)\, du| \le \sup_{|s|\le\delta} |\nu_n^x \left(g'(x,s) - g'(x,0)\right)|.$$

Stochastic equicontinuity condition for empirical process indexed by $\{g'(\cdot,t) : t \in T\}$ gives $R_n(t) = o_p(|t|)$ uniformly in shrinking neighborhoods of $t = 0$.
See Pollard (1996*a*, Chapter 13; *retrieve ASY.RatesCid.ps from WWW site*) for applications of uniform stochastic approximations in the study of asymptotics for optimization estimators.

6

(Kim & Pollard 1990, Lemma 2.6). "Let $\{Z(t) : t \in T\}$ be a Gaussian process with continuous sample paths, indexed by a σ-compact metric space. If $\operatorname{var}(Z(s) - Z(t)) \ne 0$ for all $s \ne t$ then, with probability one, no sample path can achieve its maximum at two distinct points of T." Reduce proof to convexity fact: "If Γ_0 and Γ_1 are convex functions on $\mathbb{R}$ with infimum of right-hand derivative of Γ_0 strictly greater than supremum of right-hand derivative of Γ_1, then $\Gamma_0(z) = \Gamma_1(z)$ for at most one value of z."

7

Unimodality of $S = X_1 + \ldots + X_n$, where the X_i are independent and X_i is $\text{Bin}(1, p_i)$ distributed. Show ratios $\mathbb{P}\{S = k\}/\mathbb{P}\{S = k-1\}$ decrease as k increases. That is, show

$$\mathbb{P}\{S = k+1\}\mathbb{P}\{S = k-1\} \le \mathbb{P}\{S = k\}^2.$$

✠ Independent copy $T = X_1' + \ldots + X_n'$. Condition on the $W_i = X_1 + X_i'$. Generate X_i and X_i' from W_i: toss fair coin if $W_i = 1$. If n_j of the W_i equal j (for $j = 0, 1, 2$) then number of heads H has $\text{Bin}(n_1, 1/2)$ conditional distribution. $\mathbb{P}\{S = k,\ T = k \mid \mathbf{W}\}$ is zero unless $2k = n_1 + 2n_2$, in which case conditional probability equals $\mathbb{P}\{H = n_1/2 \mid \mathbf{W}\}$, which is greater than $\mathbb{P}\{H = 1 + n_1/2 \mid \mathbf{W}\}$. For more surprising consequences see Samuels (1965) and Jogdeo & Samuels (1968).

References

Jogdeo, K. & Samuels, S. M. (1968), 'Monotone convergence of binomial probabilities and a generalization of Ramanujan's theorem', *Annals of Mathematical Statistics* **39**, 1191–1195.

Kim, J. & Pollard, D. (1990), 'Cube root asymptotics', *Annals of Statistics* **18**, 191–219.

Pollard, D. (1991), 'Asymptotics for least absolute deviation regression estimators', *Econometric Theory* **7**, 186–199.

Pollard, D. (1996*a*), *Asymptopia: An Exposition of Statistical Asymptotic Theory*, ??? (Unpublished book manuscript.).

Pollard, D. (1996*b*), *An Explanation of Probability*, ??? (Unpublished book manuscript.).

Samuels, S. M. (1965), 'On the number of successes in independent trials', *Annals of Mathematical Statistics* **36**, 1272–1278.

Shorack, G. R. & Wellner, J. A. (1986), *Empirical Processes with Applications to Statistics*, Wiley, New York.

David Pollard: July 1995, Kingston