proof of Chernoff-Cramer bound

Canonical name	ProofOfChernoffCramerBound
Date of creation	2013-03-22 16:09:05
Last modified on	2013-03-22 16:09:05
Owner	Andrea Ambrosio (7332)
Last modified by	Andrea Ambrosio (7332)
Numerical id	26
Author	Andrea Ambrosio (7332)
Entry type	Proof
Classification	msc 60E15

Let $h(x)$ be the step function ($h(x) = 1$ for $x \geq 0$, $h(x) = 0$ for $x < 0$); then, by generalized Markov inequality, for any $t > 0$ and any $\varepsilon \geq 0$,

$$\begin{aligned}
\Pr\left\{\sum_{i=1}^{n}(X_i - E[X_i]) > \varepsilon\right\} &= E\left[h\left(\sum_{i=1}^{n}(X_i - E[X_i]) - \varepsilon\right)\right] \leq \\
&\leq E\left[e^{t\left(\sum_{i=1}^{n}(X_i - E[X_i]) - \varepsilon\right)}\right] = \\
&= \exp(-\varepsilon t) E\left[e^{\sum_{i=1}^{n} t(X_i - E[X_i])}\right] = \\
&= \exp(-\varepsilon t) E\left[\prod_{i=1}^{n} e^{t(X_i - E[X_i])}\right] = \\
\text{(by independence)} &= \exp(-\varepsilon t) \prod_{i=1}^{n} E\left[e^{t(X_i - E[X_i])}\right] = \\
&= \exp\left(-\varepsilon t + \sum_{i=1}^{n} \ln E\left[e^{t(X_i - E[X_i])}\right]\right) = \\
&= \exp\left[-\left(t\varepsilon - \psi(t)\right)\right].
\end{aligned}$$

Since this expression is valid for any $t > 0$, the best bound is obtained taking the supremum:

$$\Pr\left\{\sum_{i=1}^{n}(X_i - E[X_i]) > \varepsilon\right\} \leq e^{-\sup_{t>0}(t\varepsilon - \psi(t))}$$

which proves part c).

To prove part a), let's observe that $\Psi(0) = \sup_{t>0}(-\psi(t)) = -\inf_{t>0}(\psi(t))$ and that

$$\begin{aligned}
E\left[e^{t(X_i - E[X_i])}\right] &\geq E[1 + t(X_i - E[X_i])] = \\
&= E[1] + tE[X_i] - tE[E[X_i]] = \\
&= 1 = E\left[e^{t(X_i - E[X_i])}\right]_{t=0}
\end{aligned}$$

that is, $t = 0$ is the infimum point for $E\left[e^{t(X_i - E[X_i])}\right]$ $\forall i$ and consequently for $\psi(t) = \sum_{i=1}^{n} \ln E\left[e^{t(X_i - EX_i)}\right]$, so as a conclusion $\Psi(0) = -\psi(0) = 0$

b) Let $x > 0$ be fixed and let t_0 be the supremum point for $tx - \psi(t)$; we have to show that $t_0 x - \psi(t_0) > 0$.

By differentiation, $\psi'(t_0) = x$.

Let's recall that the moment generating function is convex, so $\psi''(t) > 0$. Writing the Taylor expansion for $\psi(t)$ around $t = t_0$, we have, with a suitable $t_1 < t_0$,

$$0 = \psi(0) = \psi(t_0) - \psi'(t_0)t_0 + \frac{1}{2}\psi''(t_1)t_0^2$$

that is

$$\Psi(x) = t_0 x - \psi(t_0) = t_0\psi'(t_0) - \psi(t_0) = \frac{1}{2}\psi''(t_1)t_0^2 > 0$$

The convexity of $\Psi(x)$ follows from the fact that $\Psi(x)$ is the supremum of the linear (and hence convex) functions $tx - \psi(t)$ and so must be convex itself.

Eventually, in to prove that $\Psi(x)$ is an increasing function, let's note that

$$\Psi'(0) = \lim_{x\to 0}\frac{\Psi(x) - \Psi(0)}{x} = \lim_{x\to 0}\frac{\Psi(x)}{x} > 0$$

and that, by Taylor formula with Lagrange form remainder, for a $\xi = \xi(x)$

$$\Psi'(x) = \Psi'(0) + \Psi''(\xi)x \geq 0$$

since $\Psi''(\xi) \geq 0$ by convexity and $x \geq 0$ by hypotheses.