proof of Bennett inequality

Canonical name	ProofOfBennettInequality
Date of creation	2013-03-22 16:12:27
Last modified on	2013-03-22 16:12:27
Owner	Andrea Ambrosio (7332)
Last modified by	Andrea Ambrosio (7332)
Numerical id	22
Author	Andrea Ambrosio (7332)
Entry type	Proof
Classification	msc 60E15

By http://planetmath.org/ChernoffCramerBoundChernoff-Cramèr inequality, we have:

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

where

$$\psi(t) = \sum_{i=1}^{n}\left(\ln E\left[e^{tX_i}\right] - tE\left[X_i\right]\right).$$

Keeping in mind that the condition

$$\Pr\{|X_i| \leq M\} = 1 \quad \forall i$$

implies that, for all i,

$$E[|X_i|^k] \leq M^k \quad \forall k \geq 0$$

(see http://planetmath.org/RelationBetweenAlmostSurelyAbsolutelyBoundedRandomVariab
for a proof) and since $\ln x \leq x - 1\ \forall x > 0$, and

$$E[|X|^k] \leq M^k \quad \Longrightarrow \quad E\left[|X|^k\right] \leq E\left[X^2\right]M^{k-2} \qquad \forall k \geq 2, k \in N$$

(see http://planetmath.org/AbsoluteMomentsBoundingNecessaryAndSufficientConditionhe for a proof), one has:

$$\begin{aligned}
\psi(t) &= \sum_{i=1}^{n} \left(\ln E\left[e^{tX_i}\right] - tE\left[X_i\right] \right) \\
&\leq \sum_{i=1}^{n} E\left[e^{tX_i}\right] - tE\left[X_i\right] - 1 \\
&= \sum_{i=1}^{n} E\left[\sum_{k=0}^{\infty} \frac{(tX_i)^k}{k!} \right] - tE\left[X_i\right] - 1 \\
&= \sum_{i=1}^{n} \left(\sum_{k=0}^{\infty} \frac{t^k E\left[X_i^k\right]}{k!} \right) - tE\left[X_i\right] - 1 \\
&= \sum_{i=1}^{n} \left(\sum_{k=2}^{\infty} \frac{t^k E\left[X_i^k\right]}{k!} \right) \\
&\leq \sum_{i=1}^{n} \left(\sum_{k=2}^{\infty} \frac{t^k E\left[\left|X_i\right|^k\right]}{k!} \right) \\
&\leq \sum_{i=1}^{n} \left(\sum_{k=2}^{\infty} \frac{t^k E\left[X_i^2\right] M^{k-2}}{k!} \right) \\
&= \sum_{k=2}^{\infty} \frac{t^k M^{k-2} \sum_{i=1}^{n} E\left[X_i^2\right]}{k!} \\
&= \frac{v^2}{M^2} \sum_{k=2}^{\infty} \frac{(tM)^k}{k!} \\
&= \frac{v^2}{M^2} \left[\exp(tM) - tM - 1 \right]
\end{aligned}$$

One can now write

$$\sup_{t \geq 0} \left(t\varepsilon - \psi(t) \right) \geq \sup_{t \geq 0} \left(t\varepsilon - \frac{v^2}{M^2} \left(e^{tM} - tM - 1 \right) \right) = \sup_{t > 0} \left[\frac{v^2}{M^2} \left(\frac{M^2 \varepsilon}{v^2} t - \left(e^{tM} - tM - 1 \right) \right) \right].$$

By elementary calculus, we obtain the value of t that maximizes the expression in round brackets:

$$t_{opt} = \frac{1}{M} \ln \left(1 + \frac{M\varepsilon}{v^2} \right)$$

which, once plugged into the bound, yields

$$\Pr\left\{\sum_{i=1}^{n}\left(X_i - E[X_i]\right) > \varepsilon\right\} \leq \exp\left[-\frac{v^2}{M^2}\left(\left(1+\frac{M\varepsilon}{v^2}\right)\ln\left(1+\frac{M\varepsilon}{v^2}\right)-\frac{M\varepsilon}{v^2}\right)\right].$$

Observing that $(1+x)\ln(1+x)-x \geq \frac{x}{2}\ln(1+x)\ \forall x \geq 0$ (see `http://planetmath.org/ASimpl`
one gets the sub-optimal yet more easily manageable formula:

$$\Pr\left\{\sum_{i=1}^{n}\left(X_i - E[X_i]\right) > \varepsilon\right\} \leq \exp\left[-\frac{\varepsilon}{2M}\ln\left(1+\frac{\varepsilon M}{v^2}\right)\right].$$