MATHEMATICS

Bernstein inequality

Asked 3 years, 6 months ago Active today Viewed 287 times

4

Let us assume $X_1,...,X_n$ are independent random variables bounded by the interval $[a_i, b_i]$ and $S_n = X_1 + ... + X_n$. When $|X_i - E[X_i]| \leq M$, the Bernstein's inequality suggests the following. It can be asumed that $M = \max_i \{b_i - E[X_i]\}$.

$$P(S_n - E[S_n] > t) \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{n}\mathrm{Var}(X_i) + \frac{2}{3}Mt}\right).$$

Now, I have a case where $Y_1,...,Y_{n_1}$ are independent random variables bounded by the interval $[c_i, d_i]$ and $S_{n_1} = Y_1 + ... + Y_{n_1}$. In addition, $Z_1,...,Z_{n_2}$ are independent random variables bounded by the interval $[e_i, f_i]$ and $S_{n_2} = Z_1 + ... + Z_{n_2}$. Y_i's and Z_i's are also independent. Let, $M_1 = \max_i \{d_i - E[Y_i]\}$ and $M_2 = \max_i \{f_i - E[Z_i]\}$. I would like to have a Bernstein's bound for $P(S_{n_1} + S_{n_2} - E[S_{n_1} + S_{n_2}] > t)$.

My try:

$$\begin{aligned} &P(S_{n_1} + S_{n_2} - E[S_{n_1} + S_{n_2}] > t) \\ &\leq \exp\left(\frac{-t^2}{2\left[\sum_{i=1}^{n_1}\mathrm{Var}(Y_i) + \sum_{i=1}^{n_2}\mathrm{Var}(Z_i)\right] + \frac{2}{3}[M_1 + M_2]t}\right). \end{aligned}$$

I am wondering whether the above equation is correct.

probability inequality random-variables independence

Share Cite Edit Follow Flag

edited 5 hours ago
Lee David Chung Lin
6,667 9 23 48

asked Jun 3 '18 at 21:10
Mike Kehoe
63 4

1 Answer

Active Oldest Votes

2

I can think of at least two "direct applications" of Bernstein inequality, and they are different from yours. I wouldn't say yours is incorrect, but to me it is not a "direct application".

First Direct Application

Consider combining all Y_i and all Z_i as just one set. In short, this gives the term $\frac{2}{3}\max\{M_1, M_2\} \cdot t$ instead of $\frac{2}{3}[M_1 + M_2] \cdot t$ in your expression, with other terms all the same.

Since this is in the denominator of negative exponent, your $M_1 + M_2 > \max\{M_1, M_2\}$ is **more conservative**, with the whole $\exp(-\text{blah})$ being larger.

Formal justification of the above if needed:

Since Y_i and Z_i are independent within each set and to each other, along with the upper bounds d_i and f_i being distinct to begin with, we can combine Y_i and Z_i as just one set.

That is, we have a set for $i = 1, 2, \ldots, (n_2 + n_1)$ that shall be denoted W_i, which bounding intervals are $[c_i, d_i]$ for the first n_1 terms and $[e_{i-n_1}, f_{i-n_1}]$ for the remaining $i = 1 + n_1, 2 + n_1, \ldots, n_2 + n_1$. (the c_i, d_i, e_i, f_i are given as in your question statement)

Thus, applying the definition (quoting your statement in the question post) $M = \max_i \left\{b_i - E[X_i]\right\}$, here we have the "relevant M" as

$$\max\left\{\max_{i=1\sim n_1}\left\{d_i - E[Y_i]\right\},\ \max_{i=1\sim n_2}\left\{f_i - E[Z_i]\right\}\right\} = \max\{M_1, M_2\}$$

Second Direct Application

Consider the equivalent statement of the inequality in terms of the complement (CDF instead of the tail):

$$P\left(S_{n_1} - E[S_{n_1}] \leq x\right) > \mathcal{P}_1(x) \equiv 1 - \exp\left[-x^2\left(2\sum_{i=1}^{n_1}\operatorname{Var}(Y_i) + \frac{2}{3}M_1 x\right)^{-1}\right]$$
$$P\left(S_{n_2} - E[S_{n_2}] \leq x\right) > \mathcal{P}_2(x) \equiv 1 - \exp\left[-x^2\left(2\sum_{i=1}^{n_2}\operatorname{Var}(Z_i) + \frac{2}{3}M_2 x\right)^{-1}\right]$$

again, all the S_{n_1} etc are as defined by you.

The desired probability is a convolution-like integral, due to the direct product of probabilities from independence:

$$\begin{aligned}
&P\left(S_{n_1} + S_{n_2} - E[S_{n_1} + S_{n_2}] > t\right) \\
&= 1 - P\left(S_{n_1} + S_{n_2} - E[S_{n_1} + S_{n_2}] \leq t\right) \\
&= 1 - \int_{u=-\infty}^{\infty} P\left(S_{n_1} - E[S_{n_1}] \leq t\right) \cdot P\left(S_{n_2} - E[S_{n_2}] \leq t - u\right)\, \mathrm{d}u \\
&\leq 1 - \int_{u=-\infty}^{\infty} \mathcal{P}_1(u)\mathcal{P}_2(t-u)\, \mathrm{d}u
\end{aligned}$$

Once you figure out the proper range for t to replace the integration lower limit $-\infty$ and upper ∞, this integral is not difficult.

Anyway, this is what I consider a "direct application" of Bernstein inequality, and it's not the same as the one presented (unless there's some more steps pushing the inequality in a way I cannot imagine).

Share Cite Edit Follow Flag

edited 4 hours ago

answered Jun 6 '18 at 22:06

Lee David Chung Lin

6,667 9 23 48

Thanks for your response. I agree with the explanation. – Mike Kehoe Jun 7 '18 at 4:54