Canad. Math. Bull. Vol. **16** (1), 1973

# BERNSTEIN'S INEQUALITY IN THE BIVARIATE CASE

BY
KENNETH MULLEN

SUMMARY. If $X_1, X_2, \ldots, X_n$, is a set of $n$ independent random variables, such that $EX_i=0$, $\mathrm{Var}(X_i)=\sigma_i^2$, and if $t$ is a real positive number and $\sigma^2=\Sigma_i\sigma_i^2$, then Bernstein [2] has given an upper bound for $\Pr[\Sigma X_i\geq t\sigma]$ when the $X$'s are bounded. The best English language discussion of Bernstein's work is probably by Bennett [1]. In this paper we consider the bivariate case where random vectors $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ are observed, where $EX_i=EY_i=0$, $\mathrm{Var}(X_i)=\mathrm{Var}(Y_i)=\sigma_i^2$, $EX_iY_i=\rho\sigma_i^2$. An expression for the upper bound for $\Pr[\Sigma X_i\geq t\sigma, \Sigma Y_i\geq t\sigma]$ is given when both $X$ and $Y$ are bounded.

**Derivation.**

THEOREM. *Let $t$, $a$, and $b$ be real positive numbers and consider $n$ independent random variables $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ where $EX_i=EY_i=0$, $\mathrm{Var}(X_i)=\mathrm{Var}(Y_i)=\sigma_i^2$, $EX_iY_i=\rho\sigma_i^2$, $\sigma^2=\sum_{i=1}^n\sigma_i^2$ and for which $|X_i|\leq R$, $|Y_i|\leq R$, then*

$$P\left[\sum_i X_i\geq t\sigma, \sum_i Y_i\geq t\sigma\right] < \exp\left\{\frac{-t^2(2-|\rho|)}{2\left(1+\frac{tR}{3\sigma}\right)}\right\}.$$

**Proof.** Consider $e^{aX_i+bY_i}$, then

$$e^{aX_i+bY_i}=\left\{1+aX_i+\sum_{r=2}^{\infty}\frac{a^rX_i^r}{r!}\right\}\left\{1+bY_i+\sum_{s=2}^{\infty}\frac{b^sY_i^s}{s!}\right\}.$$

Expanding, taking expectations, and noting that $EaX_i=EbY_i=0$, we get

$$Ee^{aX_i+bY_i}=1+ab\rho\sigma_i^2+\frac{a^2\sigma_i^2A_i}{2}+\frac{b^2\sigma_i^2B_i}{2}+ab\sigma_i^2(C_i+D_i+E_i)$$

where

$$A_i=\sum_{r=2}^{\infty}\frac{a^{r-2}EX_i^r}{r!\sigma_i^2\frac{1}{2}}$$

$$B_i=\sum_{s=2}^{\infty}\frac{b^{s-2}EY_i^2}{s!\sigma_i^2\frac{1}{2}}$$

(1)
$$C_i=\sum_{r=2}^{\infty}\frac{a^{r-1}EY_iX_i^r}{r!\sigma_i^2}$$

$$D_i=\sum_{s=2}^{\infty}\frac{b^{s-1}EX_iY_i^s}{s!\sigma_i^2}$$

$$E_i=\sum_{r=2}^{\infty}\sum_{s=2}^{\infty}\frac{a^{r-1}b^{s-1}EX_i^rY_i^s}{r!s!\sigma_i^2}.$$

Received by the editors September 3, 1970 and, in revised form, August 10, 1971.



https://doi.org/10.4153/CMB-1973-017-5 Published online by Cambridge University Press

Thus we can write

$$Ee^{aX_i+bY_i} < \exp\left\{\left(\frac{a^2\sigma_i^2}{2}+\frac{b^2\sigma_i^2}{2}\right)M_i+ab\sigma_i^2G_i\right\}$$

where $M_i=\text{Max}(A_i, B_i)$,

$$G_i = \rho+C_i+D_i+E_i.$$

Now if $S(X)=\sum_{i=1}^n X_i$, $S(Y)=\sum_{i=1}^n Y_i$ then

$$Ee^{aS(X)+bS(Y)} < \exp\left\{\left(\frac{a^2\sigma^2}{2}+\frac{b^2\sigma^2}{2}\right)M+ab\sigma G\right\}$$

where $M=\max M_i$, $G=\max G_i$.

Let $h(X, Y)$ be a nonnegative function of $X$ and $Y$ with p.d.f. $f(x,y)$, so that $h(X, Y)>K$ when $X\geq C_1$, $Y\geq C_2$.

$$E[h(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)f(x, y)\,dx\,dy \geq \int_{x\geq C_1}\int_{y\geq C_2} h(x, y)f(x, y)\,dx\,dy$$

$$\geq K\int_{x\geq C_1}\int_{y\geq C_2} f(x, y)\,dx\,dy = K\cdot P[X\geq C_1, Y\geq C_2].$$

Now if $h(X, Y)=\exp(aX+bY)$ then

$$P[X\geq C_1, Y\geq C_2] \leq \frac{E\exp(aX+bY)}{\exp(aC_1+bC_2)}.$$

Replace $X$ by $S(X)$, $Y$ by $S(Y)$, $C_1=C_2=t\sigma$, then

$$P[S(X)\geq t\sigma, S(Y)\geq t\sigma] \leq \frac{E\exp(aS(X)+bS(Y))}{\exp((a+b)(t\sigma))}$$

$$< \exp\left\{\left(\frac{a^2\sigma^2}{2}+\frac{b^2\sigma^2}{2}\right)M+ab\sigma^2G-(t\sigma)(a+b)\right\}.$$

This right-hand side is minimized with respect to $a$ and $b$ for

$$a\sigma^2M+b\sigma^2G = t\sigma$$
$$b\sigma^2M+a\sigma^2G = t\sigma$$

i.e. when

(2) $$M = \frac{t}{\sigma(a+b)}, \qquad G = \frac{t}{\sigma(a+b)}.$$

Thus we can write:

(3) $$P[S(X)\geq t\sigma, S(Y)\geq t\sigma] < \exp\left\{\frac{-t\sigma(a+b)}{2}\right\}.$$

Now suppose that

$$\left.\begin{aligned} E\,|X_i|^r &\leq \tfrac{1}{2}\sigma_i^2 r!W^{r-2} \\ E\,|Y_i|^s &\leq \tfrac{1}{2}\sigma_i^2 s!W^{r-2}\end{aligned}\right\} \qquad \begin{aligned} &r, s\geq 2\\ &W \text{ a constant}\end{aligned}$$

then since $EX_i^r\leq E\,|X_i|^r$ we have

$$A_i\leq\frac{1}{1-aW}, \qquad B_i\leq\frac{1}{1-bW}.$$

https://doi.org/10.4153/CMB-1973-017-5 Published online by Cambridge University Press

Now without loss of generality assume $a>b$ ($a=b$ can be omitted), so that we can state $M_i \le (1-aW)^{-1}$, which being independent of $i$ implies that

$$M \le \frac{1}{1-aW}. \tag{4}$$

Further suppose that

$$E\,|Y_i|\,|X_i|^r \le |\rho|\,\sigma_i^2 r!\,W^{r-1}$$
$$E\,|X_i|\,|Y_i|^s \le |\rho|\,\sigma_i^2 s!\,W^{s-1}$$
$$E\,|X_i|^r\,|Y_i|^s \le |\rho|\,\sigma_i^2 r!s!\,W^{r-1}W^{s-1}$$

where $r, s \ge 2$, $W$ is a constant, then

$$C_i \le \frac{|\rho|\,aW}{1-aW}$$
$$D_i \le \frac{|\rho|\,bW}{1-bW}$$
$$E_i \le \frac{|\rho|\,abW^2}{(1-aW)(1-bW)}.$$

Noting that the right-hand sides are independent of $i$, we have

$$\begin{aligned} G &\le |\rho|\left\{1+\frac{aW}{1-aW}+\frac{bW}{1-bW}+\frac{abW^2}{(1-aW)(1-bW)}\right\} \\ &= \frac{|\rho|}{(1-aW)(1-bW)}. \end{aligned} \tag{5}$$

Let us now choose $W$ so that

$$1-bW \le |\rho| \tag{6}$$

then from (4), (5), and (6)

$$M = \frac{t}{\sigma(a+b)} \le \frac{1}{1-aW} \le \frac{|\rho|}{(1-aW)(1-bW)}$$

Taking the left most inequality yields

$$a \ge \frac{t-\sigma b}{\sigma+tW}$$

which, when put into (3) gives

$$P[S(X) \ge t\sigma, S(Y) \ge t\sigma] < \exp\left\{\frac{-t^2\sigma}{2}\,\frac{(1+bW)}{\sigma+tW}\right\}.$$

Using (6) gives

$$P[S(X) \ge t\sigma, S(Y) \ge t\sigma] < \exp\left\{\frac{-t^2(2-|\rho|)}{2(\sigma-tW)}\right\}$$

or

$$P[S(X) \ge t\sigma, S(Y) \ge t\sigma] < \exp\left\{\frac{-t^2(2-|\rho|)}{2\left(1+t\frac{W}{\sigma}\right)}\right\}. \tag{7}$$

https://doi.org/10.4153/CMB-1973-017-5 Published online by Cambridge University Press

Of particular interest are bounded variables satisfying

$$|X_i| \leq R, \qquad |Y_i| \leq R$$

for which

$$E\,|X_i|^r \leq \sigma_i^2 R^{r-2}, \qquad E\,|Y_i|^s \leq \sigma_i^2 R^{r-2}$$
$$E\,|Y_i|\,|X_i|^r \leq |\rho|\,\sigma_i^2 R^{r-1}$$
$$E\,|X_i|\,|Y_i|^s \leq |\rho|\,\sigma_i^2 R^{s-1}$$
$$E\,|X_i|^r\,|Y_i|^s \leq |\rho|\,\sigma_i^2 R^{s-1} R^{r-1}$$

so that $W=3R$ is the smallest value of $W$ that can be used. Then (7) becomes

$$P[S(X) \geq t\sigma, S(Y) \geq t\sigma] < \exp\left\{\frac{-t^2(2-|\rho|)}{2\left(1+\frac{tR}{3}\right)}\right\} \tag{8}$$

and the proof is complete. When $|\rho|=1$ and we have in effect only one random variable, then (8) reduces to the univariate case (see Reference [1]). When $|\rho|=0$, (8) is the product of two univariate cases.

## References

1. George Bennett, *Probability inequalities for the sum of independent random variables*, J. Amer. Statist. Assoc. **57** (1962), 33–45.

2. S. Bernstein, *Sur une modification de l'inéqualité de Tchebichef* (in Russian, French Summary), Ann. Sci. Inst. Sew. Ukraine Sect. Math. I, 1924.

University of Guelph,
Guelph, Ontario

https://doi.org/10.4153/CMB-1973-017-5 Published online by Cambridge University Press