On Feige's inequality

Krzysztof Oleszkiewicz

Institute of Mathematics
University of Warsaw & Polish Academy of Sciences

Toronto, 2010

Feige's inequality

Theorem (Uriel Feige, SIAM Journal on Computing, 2006): For every $\delta > 0$ there exists some $\varepsilon = \varepsilon(\delta) > 0$ such that for any positive integer n and any sequence of independent non-negative random variables $X_1, X_2, \ldots, X_n$ with $\mathbb{E}X_i \leq 1$ for $i = 1, 2, \ldots, n$ there is

$$\mathbb{P}(S \leq \mathbb{E}S + \delta) \geq \varepsilon(\delta),$$

where $S = X_1 + X_2 + \ldots + X_n$.

The theorem may be proved with $\liminf_{\delta \to 0} \varepsilon(\delta)/\delta > 0$ and, obviously, $\varepsilon(\delta)$ non-decreasing. It is easy to prove that, in general, one cannot hope for better asymptotics at zero.

Toulouse: Franck Barthe

Feige's inequality

Theorem (Uriel Feige, SIAM Journal on Computing, 2006): For every $\delta > 0$ there exists some $\varepsilon = \varepsilon(\delta) > 0$ such that for any positive integer n and any sequence of independent non-negative random variables $X_1, X_2, \ldots, X_n$ with $\mathbb{E}X_i \leq 1$ for $i = 1, 2, \ldots, n$ there is

$$\mathbb{P}(S \leq \mathbb{E}S + \delta) \geq \varepsilon(\delta),$$

where $S = X_1 + X_2 + \ldots + X_n$.

The theorem may be proved with $\liminf_{\delta \to 0} \varepsilon(\delta)/\delta > 0$ and, obviously, $\varepsilon(\delta)$ non-decreasing. It is easy to prove that, in general, one cannot hope for better asymptotics at zero.

Toulouse: Franck Barthe

Feige's inequality

Theorem (Uriel Feige, SIAM Journal on Computing, 2006):
For every $\delta > 0$ there exists some $\varepsilon = \varepsilon(\delta) > 0$ such that for any positive integer n and any sequence of independent non-negative random variables $X_1, X_2, \dots, X_n$ with $\mathbb{E}X_i \leq 1$ for $i = 1, 2, \dots, n$ there is

$$\mathbb{P}(S \leq \mathbb{E}S + \delta) \geq \varepsilon(\delta),$$

where $S = X_1 + X_2 + \ldots + X_n$.

The theorem may be proved with $\liminf_{\delta \to 0} \varepsilon(\delta)/\delta > 0$ and, obviously, $\varepsilon(\delta)$ non-decreasing. It is easy to prove that, in general, one cannot hope for better asymptotics at zero.

Toulouse: Franck Barthe

Reduction to two-point variables

Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent, non-negative random variables with $\mathbb{P}(\xi_i = x_i) = p_i$ and $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, where $x_i > y_i > 0$ and $p_i \in (0,1)$ for $i = 1, 2, \ldots, n$. Let $s_i = x_i - y_i$ (spread). Without loss of generality we may and will assume that $s_1 \geq s_2 \geq \ldots \geq s_n > 0$. We assume that $m_i = \mathbb{E}\xi_i \leq 1$ for all i's.

Note that $m_i - y_i = p_i s_i$, so that $p_i \leq 1/s_i$ for every $i \leq n$.

Note that distribution of any non-constant X_i from Feige's theorem is a mixture of distributions of two-point random variables of the above type (with mean $m_i = \mathbb{E}X_i$ fixed and parameters x_i, y_i and p_i varying). Since the quantity we want to estimate from below,

$$\mathbb{P}(S \leq \mathbb{E}S + \delta) = \Big(\bigotimes_{i=1}^{n} \mu_{X_i}\Big)\Big(\{t \in \mathbb{R}^n : \sum_{i=1}^{n} t_i \leq \delta + \sum_{i=1}^{n} m_i\}\Big),$$

is multilinear with respect to μ_{X_i}'s it suffices to prove the theorem with X_i's replaced by ξ_i's.

Reduction to two-point variables

Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent, non-negative random variables with $\mathbb{P}(\xi_i = x_i) = p_i$ and $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, where $x_i > y_i > 0$ and $p_i \in (0,1)$ for $i = 1, 2, \ldots, n$. Let $s_i = x_i - y_i$ (spread). Without loss of generality we may and will assume that $s_1 \geq s_2 \geq \ldots \geq s_n > 0$. We assume that $m_i = \mathbb{E}\xi_i \leq 1$ for all i's.

Note that $m_i - y_i = p_i s_i$, so that $p_i \leq 1/s_i$ for every $i \leq n$.

Note that distribution of any non-constant X_i from Feige's theorem is a mixture of distributions of two-point random variables of the above type (with mean $m_i = \mathbb{E}X_i$ fixed and parameters x_i, y_i and p_i varying). Since the quantity we want to estimate from below,

$$\mathbb{P}(S \leq \mathbb{E}S + \delta) = \Big(\bigotimes_{i=1}^{n} \mu_{X_i}\Big)\Big(\{t \in \mathbb{R}^n : \sum_{i=1}^{n} t_i \leq \delta + \sum_{i=1}^{n} m_i\}\Big),$$

is multilinear with respect to μ_{X_i}'s it suffices to prove the theorem with X_i's replaced by ξ_i's.

Reduction to two-point variables

Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent, non-negative random variables with $\mathbb{P}(\xi_i = x_i) = p_i$ and $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, where $x_i > y_i > 0$ and $p_i \in (0,1)$ for $i = 1, 2, \ldots, n$. Let $s_i = x_i - y_i$ (spread). Without loss of generality we may and will assume that $s_1 \geq s_2 \geq \ldots \geq s_n > 0$. We assume that $m_i = \mathbb{E}\xi_i \leq 1$ for all i's.

Note that $m_i - y_i = p_i s_i$, so that $p_i \leq 1/s_i$ for every $i \leq n$.

Note that distribution of any non-constant X_i from Feige's theorem is a mixture of distributions of two-point random variables of the above type (with mean $m_i = \mathbb{E}X_i$ fixed and parameters x_i, y_i and p_i varying). Since the quantity we want to estimate from below,

$$\mathbb{P}(S \leq \mathbb{E}S + \delta) = \Big(\bigotimes_{i=1}^{n} \mu_{X_i}\Big)(\{t \in \mathbb{R}^n : \sum_{i=1}^{n} t_i \leq \delta + \sum_{i=1}^{n} m_i\}),$$

is multilinear with respect to μ_{X_i}'s it suffices to prove the theorem with X_i's replaced by ξ_i's.

We will need the following auxiliary bound:

Proposition: For every positive C there exists $\kappa(C) > 0$ such that for any positive integer n and any sequence of independent random variables $Z_1, Z_2, \ldots, Z_n$, satisfying $\mathbb{E}Z_i = 0$ and $-K \leq Z_i \leq K$ a.s. for $i = 1, 2, \ldots, n$ and some constant K, we have

$$\mathbb{P}(Z_1 + Z_2 + \ldots + Z_n \leq C \cdot K) \geq \kappa(C).$$

Remark: In fact, much weaker assumptions suffice, for example $\mathbb{E}Z_i = 0$ and $\mathbb{E}|Z_i|^p \leq K^{p-2} \cdot \mathbb{E}Z_i^2$ for $i = 1, 2, \ldots, n$ and some fixed $p > 2$ (κ depends then both on C and p).

Auxiliary estimate

We will need the following auxiliary bound:

Proposition: For every positive C there exists $\kappa(C) > 0$ such that for any positive integer n and any sequence of independent random variables $Z_1, Z_2, \ldots, Z_n$, satisfying $\mathbb{E}Z_i = 0$ and $-K \leq Z_i \leq K$ a.s. for $i = 1, 2, \ldots, n$ and some constant K, we have

$$\mathbb{P}(Z_1 + Z_2 + \ldots + Z_n \leq C \cdot K) \geq \kappa(C).$$

Remark: In fact, much weaker assumptions suffice, for example $\mathbb{E}Z_i = 0$ and $\mathbb{E}|Z_i|^p \leq K^{p-2} \cdot \mathbb{E}Z_i^2$ for $i = 1, 2, \ldots, n$ and some fixed $p > 2$ (κ depends then both on C and p).

Proof of the auxiliary estimate

Proof of Proposition: Let $\bar{\kappa}(C)$ denote the optimal (largest) value of $\kappa(C)$ for which Proposition holds true for given $C > 0$. *A priori*, $\bar{\kappa}(C)$ may be equal to zero. Obviously, by considering symmetric ± 1 random variables and $n \to \infty$ we have $\bar{\kappa}(C) \leq 1/2$ for all $C > 0$. Also, it is clear that $\bar{\kappa}(C)$ is a non-decreasing function.

In *Concentration of capital - the product form of the LLN* (O., Stat. Probab. Letters, 2001) it is proved, with help of the Berry-Esseen inequality, that for C large enough there is $\bar{\kappa}(C) = 1/2$. Here, however, we will prove only a weaker estimate, namely:

There **exists** $C > 0$ such that $\bar{\kappa}(C) > 0$.

The amplifier trick will do the rest.

Proof of the auxiliary estimate

Proof of Proposition: Let $\bar{\kappa}(C)$ denote the optimal (largest) value of $\kappa(C)$ for which Proposition holds true for given $C > 0$. *A priori*, $\bar{\kappa}(C)$ may be equal to zero. Obviously, by considering symmetric ± 1 random variables and $n \to \infty$ we have $\bar{\kappa}(C) \leq 1/2$ for all $C > 0$. Also, it is clear that $\bar{\kappa}(C)$ is a non-decreasing function.

In *Concentration of capital - the product form of the LLN* (O., Stat. Probab. Letters, 2001) it is proved, with help of the Berry-Esseen inequality, that for C large enough there is $\bar{\kappa}(C) = 1/2$. Here, however, we will prove only a weaker estimate, namely:

There **exists** $C > 0$ such that $\bar{\kappa}(C) > 0$.

The amplifier trick will do the rest.

Proof of the auxiliary estimate

Proof of Proposition: Let $\bar{\kappa}(C)$ denote the optimal (largest) value of $\kappa(C)$ for which Proposition holds true for given $C > 0$. *A priori*, $\bar{\kappa}(C)$ may be equal to zero. Obviously, by considering symmetric ± 1 random variables and $n \to \infty$ we have $\bar{\kappa}(C) \leq 1/2$ for all $C > 0$. Also, it is clear that $\bar{\kappa}(C)$ is a non-decreasing function.

In *Concentration of capital - the product form of the LLN* (O., Stat. Probab. Letters, 2001) it is proved, with help of the Berry-Esseen inequality, that for C large enough there is $\bar{\kappa}(C) = 1/2$. Here, however, we will prove only a weaker estimate, namely:

There **exists** $C > 0$ such that $\bar{\kappa}(C) > 0$.

The amplifier trick will do the rest.

Amplifier trick

Let $S = Z_1 + Z_2 + \ldots + Z_n$. Consider i.i.d. copies of S: S_1, $S_2, \ldots$ Then

$$\mathbb{P}(S_1 + S_2 + \ldots + S_m \leq C \cdot K) \geq \bar{\kappa}(C)$$

since $S_1 + S_2 + \ldots + S_m$ is a sum of mn independent mean-zero random variables with values in $[-K, K]$ a.s.

On the other hand, we have

$$\mathbb{P}(S_1 + S_2 + \ldots + S_m \leq C \cdot K) \leq$$

$$\leq \mathbb{P}(S_1 \leq \frac{C}{m}K) + \ldots + \mathbb{P}(S_m \leq \frac{C}{m}K) = m \cdot \mathbb{P}(S \leq \frac{C}{m}K).$$

Thus we have proved that, under assumptions of Proposition, $\mathbb{P}(S \leq \frac{C}{m}K) \geq \bar{\kappa}(C)/m$, so that $\bar{\kappa}(C/m) \geq \bar{\kappa}(C)/m$ for $m \geq 1$.

Amplifier trick

Let $S = Z_1 + Z_2 + \ldots + Z_n$. Consider i.i.d. copies of S: S_1, $S_2, \ldots$ Then

$$\mathbb{P}(S_1 + S_2 + \ldots + S_m \leq C \cdot K) \geq \bar{\kappa}(C)$$

since $S_1 + S_2 + \ldots + S_m$ is a sum of mn independent mean-zero random variables with values in $[-K, K]$ a.s.

On the other hand, we have

$$\mathbb{P}(S_1 + S_2 + \ldots + S_m \leq C \cdot K) \leq$$

$$\leq \mathbb{P}(S_1 \leq \frac{C}{m}K) + \ldots + \mathbb{P}(S_m \leq \frac{C}{m}K) = m \cdot \mathbb{P}(S \leq \frac{C}{m}K).$$

Thus we have proved that, under assumptions of Proposition, $\mathbb{P}(S \leq \frac{C}{m}K) \geq \bar{\kappa}(C)/m$, so that $\bar{\kappa}(C/m) \geq \bar{\kappa}(C)/m$ for $m \geq 1$.

Amplifier trick

Let $S = Z_1 + Z_2 + \ldots + Z_n$. Consider i.i.d. copies of S: S_1, $S_2, \ldots$ Then

$$\mathbb{P}(S_1 + S_2 + \ldots + S_m \leq C \cdot K) \geq \bar{\kappa}(C)$$

since $S_1 + S_2 + \ldots + S_m$ is a sum of mn independent mean-zero random variables with values in $[-K, K]$ a.s.

On the other hand, we have

$$\mathbb{P}(S_1 + S_2 + \ldots + S_m \leq C \cdot K) \leq$$

$$\leq \mathbb{P}(S_1 \leq \frac{C}{m}K) + \ldots + \mathbb{P}(S_m \leq \frac{C}{m}K) = m \cdot \mathbb{P}(S \leq \frac{C}{m}K).$$

Thus we have proved that, under assumptions of Proposition, $\mathbb{P}(S \leq \frac{C}{m}K) \geq \bar{\kappa}(C)/m$, so that $\bar{\kappa}(C/m) \geq \bar{\kappa}(C)/m$ for $m \geq 1$.

Fourth moment bound

We have observed that it suffices to prove $\bar{\kappa}(C) > 0$ for **some** $C > 0$ to have it for **all** $C > 0$, with $\liminf_{C \to 0^+} \bar{\kappa}(C)/C > 0$.

Now, let as before $S = Z_1 + Z_2 + \ldots + Z_n$ and let $\sigma^2 = \mathbb{E}S^2$.

Note that

$$\mathbb{E}S^4 = \sum_{i=1}^{n} \mathbb{E}Z_i^4 + 6 \sum_{1 \le i < j \le n} \mathbb{E}Z_i^2 \cdot \mathbb{E}Z_j^2 \le$$

$$\le K^2 \sum_{i=1}^{n} \mathbb{E}Z_i^2 + 3\Big(\sum_{i=1}^{n} \mathbb{E}Z_i^2\Big)^2 = K^2\sigma^2 + 3\sigma^4.$$

$$\mathbb{E}S^2 = \mathbb{E}(|S|^{2/3} \cdot |S|^{4/3}) \overset{H}{\le} (\mathbb{E}|S|)^{2/3}(\mathbb{E}S^4)^{1/3}, \text{ so}$$

$$(\mathbb{E}|S|)^2/\mathbb{E}S^2 \ge \sigma^4(K^2\sigma^2 + 3\sigma^4)^{-1} = 1/(3 + K^2\sigma^{-2}).$$

Fourth moment bound

We have observed that it suffices to prove $\bar{\kappa}(C) > 0$ for **some** $C > 0$ to have it for **all** $C > 0$, with $\liminf_{C \to 0^+} \bar{\kappa}(C)/C > 0$.

Now, let as before $S = Z_1 + Z_2 + \ldots + Z_n$ and let $\sigma^2 = \mathbb{E}S^2$.

Note that

$$\mathbb{E}S^4 = \sum_{i=1}^{n} \mathbb{E}Z_i^4 + 6 \sum_{1 \le i < j \le n} \mathbb{E}Z_i^2 \cdot \mathbb{E}Z_j^2 \le$$

$$\le K^2 \sum_{i=1}^{n} \mathbb{E}Z_i^2 + 3\Big(\sum_{i=1}^{n} \mathbb{E}Z_i^2\Big)^2 = K^2\sigma^2 + 3\sigma^4.$$

$$\mathbb{E}S^2 = \mathbb{E}(|S|^{2/3} \cdot |S|^{4/3}) \overset{H}{\le} (\mathbb{E}|S|)^{2/3}(\mathbb{E}S^4)^{1/3}, \text{ so}$$

$$(\mathbb{E}|S|)^2/\mathbb{E}S^2 \ge \sigma^4(K^2\sigma^2 + 3\sigma^4)^{-1} = 1/(3 + K^2\sigma^{-2}).$$

Fourth moment bound

We have observed that it suffices to prove $\bar{\kappa}(C) > 0$ for **some** $C > 0$ to have it for **all** $C > 0$, with $\liminf_{C \to 0^+} \bar{\kappa}(C)/C > 0$.

Now, let as before $S = Z_1 + Z_2 + \ldots + Z_n$ and let $\sigma^2 = \mathbb{E}S^2$.

Note that

$$\mathbb{E}S^4 = \sum_{i=1}^{n} \mathbb{E}Z_i^4 + 6 \sum_{1 \le i < j \le n} \mathbb{E}Z_i^2 \cdot \mathbb{E}Z_j^2 \le$$

$$\le K^2 \sum_{i=1}^{n} \mathbb{E}Z_i^2 + 3(\sum_{i=1}^{n} \mathbb{E}Z_i^2)^2 = K^2\sigma^2 + 3\sigma^4.$$

$\mathbb{E}S^2 = \mathbb{E}(|S|^{2/3} \cdot |S|^{4/3}) \overset{H}{\le} (\mathbb{E}|S|)^{2/3}(\mathbb{E}S^4)^{1/3}$, so $(\mathbb{E}|S|)^2/\mathbb{E}S^2 \ge \sigma^4(K^2\sigma^2 + 3\sigma^4)^{-1} = 1/(3 + K^2\sigma^{-2})$.

Fourth moment bound

We have observed that it suffices to prove $\bar{\kappa}(C) > 0$ for **some** $C > 0$ to have it for **all** $C > 0$, with $\liminf_{C \to 0^+} \bar{\kappa}(C)/C > 0$.

Now, let as before $S = Z_1 + Z_2 + \ldots + Z_n$ and let $\sigma^2 = \mathbb{E}S^2$.

Note that

$$\mathbb{E}S^4 = \sum_{i=1}^{n} \mathbb{E}Z_i^4 + 6 \sum_{1 \le i < j \le n} \mathbb{E}Z_i^2 \cdot \mathbb{E}Z_j^2 \le$$

$$\le K^2 \sum_{i=1}^{n} \mathbb{E}Z_i^2 + 3(\sum_{i=1}^{n} \mathbb{E}Z_i^2)^2 = K^2\sigma^2 + 3\sigma^4.$$

$$\mathbb{E}S^2 = \mathbb{E}(|S|^{2/3} \cdot |S|^{4/3}) \overset{H}{\le} (\mathbb{E}|S|)^{2/3}(\mathbb{E}S^4)^{1/3}, \text{ so}$$
$$(\mathbb{E}|S|)^2/\mathbb{E}S^2 \ge \sigma^4(K^2\sigma^2 + 3\sigma^4)^{-1} = 1/(3 + K^2\sigma^{-2}).$$

Paley-Zygmund inequality

We have proved that $(\mathbb{E}|S|)^2/\mathbb{E}S^2 \geq 1/(3+K^2\sigma^{-2})$.

The classical Paley-Zygmund estimate states that

$$\mathbb{E}|S|/2 = \mathbb{E}|S|1_{S<0} \leq (\mathbb{E}S^2)^{1/2} \cdot \Big(\mathbb{P}(S<0)\Big)^{1/2},$$

so that

$$\mathbb{P}(S \leq C \cdot K) \geq \mathbb{P}(S<0) \geq \frac{(\mathbb{E}|S|)^2}{4\,\mathbb{E}S^2} \geq \frac{1}{4(3+K^2\sigma^{-2})}.$$

Thus $\mathbb{P}(S \leq C \cdot K) \geq 1/16$ if only $\sigma \geq K$,
whereas for $\sigma \leq K$ by Chebyshev's inequality we get

$$\mathbb{P}(S > C \cdot K) \leq \frac{\sigma^2}{C^2K^2} \leq C^{-2},$$

so in particular $\mathbb{P}(S \leq 2K) \geq 1 - 2^{-2} = 3/4$.

We have proved that $\overline{\kappa}(2) \geq 1/16 > 0$.

Paley-Zygmund inequality

We have proved that $(\mathbb{E}|S|)^2/\mathbb{E}S^2 \geq 1/(3+K^2\sigma^{-2})$.

The classical Paley-Zygmund estimate states that

$$\mathbb{E}|S|/2 = \mathbb{E}|S|1_{S<0} \leq (\mathbb{E}S^2)^{1/2} \cdot \Big(\mathbb{P}(S<0)\Big)^{1/2},$$

so that

$$\mathbb{P}(S \leq C \cdot K) \geq \mathbb{P}(S<0) \geq \frac{(\mathbb{E}|S|)^2}{4\,\mathbb{E}S^2} \geq \frac{1}{4(3+K^2\sigma^{-2})}.$$

Thus $\mathbb{P}(S \leq C \cdot K) \geq 1/16$ if only $\sigma \geq K$,
whereas for $\sigma \leq K$ by Chebyshev's inequality we get

$$\mathbb{P}(S > C \cdot K) \leq \frac{\sigma^2}{C^2K^2} \leq C^{-2},$$

so in particular $\mathbb{P}(S \leq 2K) \geq 1 - 2^{-2} = 3/4$.

We have proved that $\bar{\kappa}(2) \geq 1/16 > 0$.

Paley-Zygmund inequality

We have proved that $(\mathbb{E}|S|)^2/\mathbb{E}S^2 \geq 1/(3+K^2\sigma^{-2})$.

The classical Paley-Zygmund estimate states that

$$\mathbb{E}|S|/2 = \mathbb{E}|S|1_{S<0} \leq (\mathbb{E}S^2)^{1/2} \cdot \Big(\mathbb{P}(S<0)\Big)^{1/2},$$

so that

$$\mathbb{P}(S \leq C \cdot K) \geq \mathbb{P}(S<0) \geq \frac{(\mathbb{E}|S|)^2}{4\,\mathbb{E}S^2} \geq \frac{1}{4(3+K^2\sigma^{-2})}.$$

Thus $\mathbb{P}(S \leq C \cdot K) \geq 1/16$ if only $\sigma \geq K$,
whereas for $\sigma \leq K$ by Chebyshev's inequality we get

$$\mathbb{P}(S > C \cdot K) \leq \frac{\sigma^2}{C^2K^2} \leq C^{-2},$$

so in particular $\mathbb{P}(S \leq 2K) \geq 1 - 2^{-2} = 3/4$.

We have proved that $\overline{\kappa}(2) \geq 1/16 > 0$.

Paley-Zygmund inequality

We have proved that $(\mathbb{E}|S|)^2/\mathbb{E}S^2 \geq 1/(3+K^2\sigma^{-2})$.

The classical Paley-Zygmund estimate states that

$$\mathbb{E}|S|/2 = \mathbb{E}|S|1_{S<0} \leq (\mathbb{E}S^2)^{1/2} \cdot \Big(\mathbb{P}(S<0)\Big)^{1/2},$$

so that

$$\mathbb{P}(S \leq C \cdot K) \geq \mathbb{P}(S<0) \geq \frac{(\mathbb{E}|S|)^2}{4\,\mathbb{E}S^2} \geq \frac{1}{4(3+K^2\sigma^{-2})}.$$

Thus $\mathbb{P}(S \leq C \cdot K) \geq 1/16$ if only $\sigma \geq K$,
whereas for $\sigma \leq K$ by Chebyshev's inequality we get

$$\mathbb{P}(S > C \cdot K) \leq \frac{\sigma^2}{C^2K^2} \leq C^{-2},$$

so in particular $\mathbb{P}(S \leq 2K) \geq 1 - 2^{-2} = 3/4$.

We have proved that $\overline{\kappa}(2) \geq 1/16 > 0$.

Paley-Zygmund inequality

We have proved that $(\mathbb{E}|S|)^2/\mathbb{E}S^2 \geq 1/(3+K^2\sigma^{-2})$.

The classical Paley-Zygmund estimate states that

$$\mathbb{E}|S|/2 = \mathbb{E}|S|1_{S<0} \leq (\mathbb{E}S^2)^{1/2} \cdot \Big(\mathbb{P}(S<0)\Big)^{1/2},$$

so that

$$\mathbb{P}(S \leq C\cdot K) \geq \mathbb{P}(S<0) \geq \frac{(\mathbb{E}|S|)^2}{4\,\mathbb{E}S^2} \geq \frac{1}{4(3+K^2\sigma^{-2})}.$$

Thus $\mathbb{P}(S \leq C\cdot K) \geq 1/16$ if only $\sigma \geq K$,
whereas for $\sigma \leq K$ by Chebyshev's inequality we get

$$\mathbb{P}(S > C\cdot K) \leq \frac{\sigma^2}{C^2K^2} \leq C^{-2},$$

so in particular $\mathbb{P}(S \leq 2K) \geq 1-2^{-2} = 3/4$.

We have proved that $\bar{\kappa}(2) \geq 1/16 > 0$.

Proof of Feige's inequality

Recall: $\xi_1, \xi_2, \ldots, \xi_n$ are independent with $\mathbb{P}(\xi_i = x_i) = p_i$, $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, $x_i > y_i > 0$, $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread), $s_1 \geq s_2 \geq \ldots \geq s_n > 0$, $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$.

We are to prove that

$$\mathbb{P}\Big(\xi_1 + \xi_2 + \ldots + \xi_n \leq \mathbb{E}(\xi_1 + \xi_2 + \ldots + \xi_n) + \delta\Big) \geq \varepsilon(\delta).$$

Proof of Theorem: Let k be the least index i such that $p_1 s_1 + \ldots + p_i s_i \geq s_{i+1}/2$. So, $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$ but $p_1 s_1 + \ldots + p_{k-1} s_{k-1} < s_k/2$ and hence $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$.
Case 2: $s_k > 2$ and thus $p_k < 1/2$.
He, Zhang and Zhang, Math. Operations Research, 2010

Proof of Feige's inequality

Recall: $\xi_1, \xi_2, \ldots, \xi_n$ are independent with $\mathbb{P}(\xi_i = x_i) = p_i$, $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, $x_i > y_i > 0$, $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread), $s_1 \geq s_2 \geq \ldots \geq s_n > 0$, $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$.

We are to prove that

$$\mathbb{P}\Big(\xi_1 + \xi_2 + \ldots + \xi_n \leq \mathbb{E}(\xi_1 + \xi_2 + \ldots + \xi_n) + \delta\Big) \geq \varepsilon(\delta).$$

Proof of Theorem: Let k be the least index i such that $p_1 s_1 + \ldots + p_i s_i \geq s_{i+1}/2$. So, $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$ but $p_1 s_1 + \ldots + p_{k-1} s_{k-1} < s_k/2$ and hence $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$.
Case 2: $s_k > 2$ and thus $p_k < 1/2$.
He, Zhang and Zhang, Math. Operations Research, 2010

Proof of Feige's inequality

Recall: $\xi_1, \xi_2, \ldots, \xi_n$ are independent with $\mathbb{P}(\xi_i = x_i) = p_i$, $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, $x_i > y_i > 0$, $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread), $s_1 \geq s_2 \geq \ldots \geq s_n > 0$, $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$.

We are to prove that

$$\mathbb{P}\Big(\xi_1 + \xi_2 + \ldots + \xi_n \leq \mathbb{E}(\xi_1 + \xi_2 + \ldots + \xi_n) + \delta\Big) \geq \varepsilon(\delta).$$

Proof of Theorem: Let k be the least index i such that $p_1 s_1 + \ldots + p_i s_i \geq s_{i+1}/2$. So, $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$ but $p_1 s_1 + \ldots + p_{k-1} s_{k-1} < s_k/2$ and hence $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$.
Case 2: $s_k > 2$ and thus $p_k < 1/2$.
He, Zhang and Zhang, Math. Operations Research, 2010

Proof of Feige's inequality

Recall: $\xi_1, \xi_2, \ldots, \xi_n$ are independent with $\mathbb{P}(\xi_i = x_i) = p_i$, $\mathbb{P}(\xi_i = y_i) = 1 - p_i$, $x_i > y_i > 0$, $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread), $s_1 \geq s_2 \geq \ldots \geq s_n > 0$, $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$.

We are to prove that

$$\mathbb{P}\Big(\xi_1 + \xi_2 + \ldots + \xi_n \leq \mathbb{E}(\xi_1 + \xi_2 + \ldots + \xi_n) + \delta\Big) \geq \varepsilon(\delta).$$

Proof of Theorem: Let k be the least index i such that $p_1 s_1 + \ldots + p_i s_i \geq s_{i+1}/2$. So, $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$ but $p_1 s_1 + \ldots + p_{k-1} s_{k-1} < s_k/2$ and hence $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$.
Case 2: $s_k > 2$ and thus $p_k < 1/2$.
He, Zhang and Zhang, Math. Operations Research, 2010

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$; $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$; $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$, so that $\xi_k - m_k, \ldots, \xi_n - m_n$ are independent mean-zero random variables with values in $[-2, 2]$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}(\xi_1 = y_1, \ldots, \xi_{k-1} = y_{k-1}, \xi_k + \ldots + \xi_n \leq m_k + \ldots + m_n + \delta) =$$

$$(1 - p_1) \ldots (1 - p_{k-1}) \mathbb{P}\Big((\xi_k - m_k) + \ldots + (\xi_n - m_n) \leq \delta\Big) \geq$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big) \kappa(\delta/2) \geq \frac{1}{2} \kappa(\delta/2),$$

where we have used Proposition for $C = \delta/2$ and $K = 2$.

Case 1

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$;
$p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$;
$m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$, so that $\xi_k - m_k, \ldots, \xi_n - m_n$ are independent mean-zero random variables with values in $[-2, 2]$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}(\xi_1 = y_1, \ldots, \xi_{k-1} = y_{k-1}, \xi_k + \ldots + \xi_n \leq m_k + \ldots + m_n + \delta) =$$

$$(1 - p_1) \ldots (1 - p_{k-1}) \mathbb{P}\Big((\xi_k - m_k) + \ldots + (\xi_n - m_n) \leq \delta\Big) \geq$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big) \kappa(\delta/2) \geq \frac{1}{2} \kappa(\delta/2),$$

where we have used Proposition for $C = \delta/2$ and $K = 2$.

Case 1

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$; $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$; $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$, so that $\xi_k - m_k, \ldots, \xi_n - m_n$ are independent mean-zero random variables with values in $[-2, 2]$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}(\xi_1 = y_1, \ldots, \xi_{k-1} = y_{k-1}, \xi_k + \ldots + \xi_n \leq m_k + \ldots + m_n + \delta) =$$

$$(1 - p_1) \ldots (1 - p_{k-1}) \mathbb{P}\Big((\xi_k - m_k) + \ldots + (\xi_n - m_n) \leq \delta\Big) \geq$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big)\kappa(\delta/2) \geq \frac{1}{2}\kappa(\delta/2),$$

where we have used Proposition for $C = \delta/2$ and $K = 2$.

Case 1

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$; $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$; $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$.

Case 1: $s_k \leq 2$ and thus also $s_{k+1}, \ldots, s_n \leq 2$, so that $\xi_k - m_k, \ldots, \xi_n - m_n$ are independent mean-zero random variables with values in $[-2, 2]$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}(\xi_1 = y_1, \ldots, \xi_{k-1} = y_{k-1}, \xi_k + \ldots + \xi_n \leq m_k + \ldots + m_n + \delta) =$$

$$(1 - p_1) \ldots (1 - p_{k-1}) \mathbb{P}\Big((\xi_k - m_k) + \ldots + (\xi_n - m_n) \leq \delta\Big) \geq$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big)\kappa(\delta/2) \geq \frac{1}{2}\kappa(\delta/2),$$

where we have used Proposition for $C = \delta/2$ and $K = 2$.

Case 2

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$; $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$; $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$; $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$.

Case 2: $s_k > 2$ and thus $p_k < 1/2$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}\Big(\xi_1 = y_1, \ldots, \xi_k = y_k, \xi_{k+1} + \ldots + \xi_n \leq$$

$$\leq (m_1 - y_1) + \ldots + (m_k - y_k) + m_{k+1} + \ldots + m_n\Big) =$$

$$(1-p_1)\ldots(1-p_{k-1})(1-p_k)\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \leq p_1 s_1 + \ldots + p_k s_k\Big) =$$

Case 2

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$; $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$; $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$; $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$.

Case 2: $s_k > 2$ and thus $p_k < 1/2$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}\Big(\xi_1 = y_1, \ldots, \xi_k = y_k, \xi_{k+1} + \ldots + \xi_n \leq$$

$$\leq (m_1 - y_1) + \ldots + (m_k - y_k) + m_{k+1} + \ldots + m_n\Big) =$$

$$(1 - p_1)\ldots(1 - p_{k-1})(1 - p_k)\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \leq p_1 s_1 + \ldots + p_k s_k\Big) =$$

Case 2

Recall: $\mathbb{P}(\xi_i = x_i) = p_i$; $\mathbb{P}(\xi_i = y_i) = 1 - p_i$; $x_i > y_i > 0$; $p_i \in (0,1)$, $s_i = x_i - y_i$ (spread); $s_1 \geq s_2 \geq \ldots \geq s_n > 0$; $m_i = \mathbb{E}\xi_i \leq 1$, so that $p_i \leq 1/s_i$; $p_1 + \ldots + p_{k-1} < 1/2$; $p_1 s_1 + \ldots + p_k s_k \geq s_{k+1}/2$.

Case 2: $s_k > 2$ and thus $p_k < 1/2$.

$$\mathbb{P}(\xi_1 + \ldots + \xi_n \leq m_1 + \ldots + m_n + \delta) \geq$$

$$\mathbb{P}\Big(\xi_1 = y_1, \ldots, \xi_k = y_k, \xi_{k+1} + \ldots + \xi_n \leq$$

$$\leq (m_1 - y_1) + \ldots + (m_k - y_k) + m_{k+1} + \ldots + m_n\Big) =$$

$$(1 - p_1) \ldots (1 - p_{k-1})(1 - p_k) \times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \leq p_1 s_1 + \ldots + p_k s_k\Big) =$$

Case 2 ($s_k > 2$, so that $p_k < 1/2$) - the end

$$\ldots = (1 - p_1)\ldots(1 - p_{k-1})(1 - p_k)\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \le p_1 s_1 + \ldots + p_k s_k\Big) \ge$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big) \times \frac{1}{2}\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \le s_{k+1}/2\Big) \ge$$

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\kappa(1/2) = \kappa(1/2)/4,$$

where we have used Proposition for $C = 1/2$ and $K = s_{k+1}$.

Putting together both cases we finish the proof of Theorem with $\varepsilon(\delta) = \min\Big(\kappa(\delta/2)/2, \kappa(1/2)/4\Big)$.

Case 2 ($s_k > 2$, so that $p_k < 1/2$) - the end

$$\ldots = (1 - p_1)\ldots(1 - p_{k-1})(1 - p_k)\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \le p_1 s_1 + \ldots + p_k s_k\Big) \ge$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big) \times \frac{1}{2}\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \le s_{k+1}/2\Big) \ge$$

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\kappa(1/2) = \kappa(1/2)/4,$$

where we have used Proposition for $C = 1/2$ and $K = s_{k+1}$.

Putting together both cases we finish the proof of Theorem with $\varepsilon(\delta) = \min\Big(\kappa(\delta/2)/2, \kappa(1/2)/4\Big)$.

Case 2 ($s_k > 2$, so that $p_k < 1/2$) - the end

$$\ldots = (1-p_1)\ldots(1-p_{k-1})(1-p_k)\times$$

$$\mathbb{P}\Big((\xi_{k+1}-m_{k+1})+\ldots+(\xi_n-m_n)\le p_1s_1+\ldots+p_ks_k\Big)\ge$$

$$\Big(1-(p_1+\ldots+p_{k-1})\Big)\times\frac{1}{2}\times$$

$$\mathbb{P}\Big((\xi_{k+1}-m_{k+1})+\ldots+(\xi_n-m_n)\le s_{k+1}/2\Big)\ge$$

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\kappa(1/2)=\kappa(1/2)/4,$$

where we have used Proposition for $C = 1/2$ and $K = s_{k+1}$.

Putting together both cases we finish the proof of Theorem with $\varepsilon(\delta) = \min\Big(\kappa(\delta/2)/2, \kappa(1/2)/4\Big)$.

Case 2 ($s_k > 2$, so that $p_k < 1/2$) - the end

$$\ldots = (1 - p_1)\ldots(1 - p_{k-1})(1 - p_k)\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \le p_1 s_1 + \ldots + p_k s_k\Big) \ge$$

$$\Big(1 - (p_1 + \ldots + p_{k-1})\Big) \times \frac{1}{2}\times$$

$$\mathbb{P}\Big((\xi_{k+1} - m_{k+1}) + \ldots + (\xi_n - m_n) \le s_{k+1}/2\Big) \ge$$

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\kappa(1/2) = \kappa(1/2)/4,$$

where we have used Proposition for $C = 1/2$ and $K = s_{k+1}$.

Putting together both cases we finish the proof of Theorem with $\varepsilon(\delta) = \min\Big(\kappa(\delta/2)/2, \kappa(1/2)/4\Big)$.

Extension

Theorem 2: Let $t_0, M > 0$. Assume that $X_1, X_2, \ldots, X_n$ are independent random variables with $\mathbb{E}X_i = 0$ for $i = 1, 2, \ldots, n$. Assume also that they satisfy the following condition:

$$\forall_{t>t_0} \ \ \mathbb{E}X_i 1_{X_i \geq t} \geq \mathbb{E}|X_i| 1_{X_i \leq -Mt}.$$

Then for every $\delta > 0$ we have

$$\mathbb{P}(X_1 + \ldots + X_n) \geq \varepsilon(\delta, t_0, M) > 0.$$

Feige's theorem follows immediately from the case $M = 1$, $t_0 = 1$ (after centering procedure to switch from non-negative setting to mean-zero framework).

The proof goes basically along the same lines.

Extension

Theorem 2: Let $t_0, M > 0$. Assume that $X_1, X_2, \ldots, X_n$ are independent random variables with $\mathbb{E}X_i = 0$ for $i = 1, 2, \ldots, n$. Assume also that they satisfy the following condition:

$$\forall_{t>t_0} \;\; \mathbb{E}X_i 1_{X_i \geq t} \geq \mathbb{E}|X_i| 1_{X_i \leq -Mt}.$$

Then for every $\delta > 0$ we have

$$\mathbb{P}(X_1 + \ldots + X_n) \geq \varepsilon(\delta, t_0, M) > 0.$$

Feige's theorem follows immediately from the case $M = 1$, $t_0 = 1$ (after centering procedure to switch from non-negative setting to mean-zero framework).

The proof goes basically along the same lines.

Extension

Theorem 2: Let $t_0, M > 0$. Assume that $X_1, X_2, \ldots, X_n$ are independent random variables with $\mathbb{E}X_i = 0$ for $i = 1, 2, \ldots, n$. Assume also that they satisfy the following condition:

$$\forall_{t>t_0} \;\; \mathbb{E}X_i 1_{X_i \geq t} \geq \mathbb{E}|X_i| 1_{X_i \leq -Mt}.$$

Then for every $\delta > 0$ we have

$$\mathbb{P}(X_1 + \ldots + X_n) \geq \varepsilon(\delta, t_0, M) > 0.$$

Feige's theorem follows immediately from the case $M = 1$, $t_0 = 1$ (after centering procedure to switch from non-negative setting to mean-zero framework).

The proof goes basically along the same lines.