On inequalities for probabilities of unions of events and the Borel–Cantelli lemma.

Andrei Frolov
St.Petersburg State University

June 8, 2017

Absract

We discuss a method which yields upper and lower bounds for probabilities of unions of events. These bounds are stronger than the Chung–Erdős inequality and its generalizations. Relationships with Hölder's inequality and bounds for the first jumps of distribution functions are considered. The method is also applied to derive upper and lower bounds for probabilities that at least *r* from *n* events occur.

All inequalities have conditional (given σ-fields) variants which may give stronger bounds for probabilities of unions.

We discuss generalizations of the Borel–Cantelli lemma. Earlier generalizations are special cases. The conditional Borel–Cantelli lemma is mentioned as well.

1. Introduction

Inequalities for probabilities of unions of events:
Chung and Erdős (1952), Gallot (1966), Dawson and Sankoff (1967), Kounias (1968), Kwerel (1975 a,b,c), Móri and Székely (1985), Boros and Prékopa (1989), Galambos and Simonelli (1996), de Caen (1997), Kuai et al. (2000), Feng et al. (2009), Prékopa (2009), Frolov (2012, 2014, 2015a,b) and references therein. (See also references on the Borel–Cantelli lemma.)
Various methods have been applied to derive these inequalities. One of them was developed in Frolov (2012, 2015a).

2. Two examples

Let $\{A_i\}_{i=1}^n$ be events and I_{A_i} be the indicators of A_i. Put

$$U_n = \bigcup_{i=1}^n A_i, \quad \xi_n = \sum_{i=1}^n I_{A_i}, \quad p_i = \mathbf{P}(\xi_n = i),\, 0 \leqslant i \leqslant n.$$

Consider bounds for $\mathbf{P}(U_n)$ in terms of $\mu_k = \mathbf{E}\xi_n^k = \sum_{i=1}^n i^k p_i$.
Simplest bounds for $\mathbf{P}(U_n)$ are based on μ_1 and μ_2.
Fix $m \in \mathbb{N}$ with $2 \leqslant m \leqslant n$ and put $c_i = \left(1 - \frac{i}{m-1}\right)\left(1 - \frac{i}{m}\right)$ for $1 \leqslant i \leqslant n$. Then all $c_i \geqslant 0$ and

$$0 \leqslant \sum_{i=1}^n c_i p_i = \sum_{i=1}^n p_i - \frac{2m-1}{m(m-1)}\mu_1 + \frac{1}{m(m-1)}\mu_2. \tag{1}$$

Hence,

$$\mathbf{P}(U_n) = \sum_{i=1}^n p_i \geqslant \frac{2m-1}{m(m-1)}\mu_1 - \frac{1}{m(m-1)}\mu_2. \tag{2}$$

Inequality (2) holds for all m. Choose the best m.

2. Two examples

The inequality in (1) turns to equality for distributions of ξ_n concentrated in 0, $m-1$ and m. Take such distribution with two first moments μ_1 and μ_2. To this goal, we solve the system:

$$(m-1)p^*_{m-1} + mp^*_m = \mu_1,$$
$$(m-1)^2 p^*_{m-1} + m^2 p^*_m = \mu_2.$$

Then

$$p^*_{m-1} = \frac{m\mu_1 - \mu_2}{m-1}, \quad p^*_m = \frac{\mu_2 - (m-1)\mu_1}{m}.$$

Since $p^*_{m-1} \geqslant 0$ and $p^*_m \geqslant 0$, we have $\mu_2/\mu_1 \leqslant m \leqslant \mu_2/\mu_1 + 1$. Putting $m = \mu_2/\mu_1 - \theta + 1$ in (2), we get the Dawson–Sankoff inequality:

$$\mathbf{P}(U_n) \geqslant \frac{\theta\mu_1^2}{\mu_2 + (1-\theta)\mu_1} + \frac{(1-\theta)\mu_1^2}{\mu_2 - \theta\mu_1}, \tag{3}$$

where $\theta = \mu_2/\mu_1 - [\mu_2/\mu_1]$ and $[x]$ is the integer part of x. Note that θ may be positive. The right-hand side of (3) is minimal for $\theta = 0$. It yields the Chung–Erdős inequality: $\mathbf{P}(U_n) \geqslant \frac{\mu_1^2}{\mu_2}$.

2. Two examples

Instead of ξ_n, consider the r.v. η_n with $q_i = \mathbf{P}(\eta_n = x_i)$ for $1 \leqslant i \leqslant n$, and $0 < x_1 < x_2 < \cdots < x_n$. Put $s_1 = \mathbf{E}\eta_n = \sum_{i=1}^n x_i q_i$ and $s_2 = \mathbf{E}\eta_n^2 = \sum_{i=1}^n x_i^2 q_i$. An analogue of (1) is

$$0 \leqslant \sum_{i=1}^n \left(1 - \frac{x_i}{x_{m-1}}\right)\left(1 - \frac{x_i}{x_m}\right) q_i = 1 - \frac{x_m + x_{m-1}}{x_m x_{m-1}} s_1 + \frac{1}{x_m x_{m-1}} s_2. \quad (4)$$

For positive $\{\alpha_i\}_{i=1}^n$ and $\{\beta_i\}_{i=1}^n$ with $\sum_{i=1}^n \alpha_i^2 = \sum_{i=1}^n \beta_i^2 = 1$, put $q_i = \beta_i^2$ and $x_i = \alpha_i/\beta_i$ for $1 \leqslant i \leqslant n$. (Assume now for simplicity that all $x_i < x_{i+1}$.) Then (4) turns to

$$0 \leqslant 1 - \frac{x_m + x_{m-1}}{x_m x_{m-1}} \sum_{i=1}^n \alpha_i \beta_i + \frac{1}{x_m x_{m-1}}.$$

This gives an upper bound for $\sum_{i=1}^n \alpha_i \beta_i$ which is $\leqslant 1$ for some m^*. For the Chung–Erdős inequality, we have the analogue:

$$\sum_{i=1}^n \alpha_i \beta_i \leqslant 1.$$

This is the Cauchy–Bunyakovski inequality.

3. Bounds for numbers.

All vectors are columns. For every $\mathbf{v} \in \mathbb{R}^d$, let v_j, $1 \leqslant j \leqslant d$ be its coordinates. We write $\mathbf{v} \leqslant \mathbf{u}$ for $\mathbf{v}, \mathbf{u} \in \mathbb{R}^d$, if $v_j \leqslant u_j$ for all j. Put $\mathbf{0}_d = (0, 0, \dots, 0)^T \in \mathbb{R}^d$ and $\mathbf{1}_d = (1, 1, \dots, 1)^T \in \mathbb{R}^d$.

Theorem (1)

Let $\mathbf{r} \in \mathbb{R}^n$ be a vector with $\mathbf{r} \geqslant \mathbf{0}_n$ and $\mathbf{F} = (f_{ki})_{k=1,i=1}^{\ell,n}$ be a matrix with real entries, where $2 \leqslant \ell \leqslant n$. Put $R = \sum\limits_{i=1}^{n} r_i$ and

$$\bar{\mathbf{s}} = \mathbf{F}\mathbf{r}. \tag{5}$$

Assume that for some $\mathbf{i} \in \mathbb{R}^\ell$ such that $1 \leqslant i_1 < i_2 < \dots < i_\ell \leqslant n$, the vector $\mathbf{a} \in \mathbb{R}^\ell$ is the solution of

$$\mathbf{F}_{\mathbf{i}}^T \mathbf{a} = \mathbf{1}_\ell, \tag{6}$$

where $\mathbf{F}_{\mathbf{i}} = (f_{ki_j})_{k=1,j=1}^{\ell,\ell}$. Put

$$\mathbf{c} = \mathbf{1}_n - \mathbf{F}^T \mathbf{a}. \tag{7}$$

3. Bounds for numbers.

Let $\mathbf{r}^* \in \mathbb{R}^n$ *be a vector such that* $\mathbf{r}^*_{\mathbf{i}} = (r^*_{i_1}, r^*_{i_2}, \ldots, r^*_{i_\ell})^T$ *is the solution of the system*

$$\mathbf{F}_{\mathbf{i}} \mathbf{r}^*_{\mathbf{i}} = \bar{\mathbf{s}} \tag{8}$$

and $r^*_i = 0$ *for all* $i \neq i_k$, $1 \leqslant i \leqslant n$. *Put* $R^* = \sum_{i=1}^{n} r^*_i$.
If $\mathbf{c} \geqslant \mathbf{0}_n$, *then* $R \geqslant R^*$. *If* $\mathbf{c} \leqslant \mathbf{0}_n$, *then* $R \leqslant R^*$. $\square$

In the examples: $\ell = 2$, $r_i = p_i$, $f_{ki} = i^k$, $i_1 = m - 1$, $i_2 = m$, $c_i = (1 - i/(m-1))(1 - i/m) \geqslant 0$, a_1 and a_2 – coefficients at μ_1 and μ_2, $r_{i_j} = p^*_{i_j}$. For the second example $f_{ki} = x_i^k$.
Theorem 1 yields lower bounds. Finally, an optimization over m.

The way of application of Theorem 1. Take ℓ, $\mathbf{F} = (f_{ki})$ (above choice only), $\mathbf{i}$. Solve (6) (or (8) for invertible $\mathbf{F}_{\mathbf{i}}$). Find $\mathbf{c}$ and check $\mathbf{c} \geqslant \mathbf{0}_n$ or $\mathbf{c} \leqslant \mathbf{0}_n$. For $\mathbf{i} = \mathbf{i}(m)$, make an optimization.

3. Bounds for numbers.

Choice of $\mathbf{i}$ comes from the special case

$$c_i = \prod_{j=1}^{\ell}(1 - \frac{i}{i_j}).$$

Lower bounds ($c_i \geqslant 0$).
$\ell = 2$: $\mathbf{i} = (m-1, m)$,
$\ell = 3$: $\mathbf{i} = (m-1, m, n)$,
$\ell = 4$: $\mathbf{i} = (m-3, m-2, m-1, m)$ and so on.

Upper bounds ($c_i \leqslant 0$). $\ell = 2$: $\mathbf{i} = (1, n)$, $\ell = 3$: $\mathbf{i} = (1, m-1, m)$,
$\ell = 4$: $\mathbf{i} = (1, m-2, m-1, n)$ and so on.

3. Bounds for numbers.

Properties of bounds:
1. they are sharp, i.e. inequalities may turn to equalities. Assume that for R, we can construct R^*. Put $R = R^*$. Bounds for such R is R^* as well.
2. Let $R^*(\ell)$ and $R^*(\ell-1)$ be lower bounds based on $(\bar{s}_1, \dots, \bar{s}_\ell)$ and $(\bar{s}_1, \dots, \bar{s}_{\ell-1})$. Then $R^*(\ell) \geqslant R^*(\ell-1)$.
This follows from Theorem 1 since for $R = R^*(\ell)$ we get $R^* = R^*(\ell-1)$ using $(\bar{s}_1, \dots, \bar{s}_{\ell-1})$.
Similarly, we have an opposite inequality for lower bounds.
So, if the number of moments increases then bounds improve.

4. Lower bounds for $\mathbf{P}(U_n)$.

Simplest variant $r_i = p_i$ in Theorem 1.

Theorem (2)

Let $\ell = 2$, $0 < a < b$, $f_{1i} = i^a$ *and* $f_{2i} = i^b$ *for all* i. *Put*
$\bar{\delta} = (\bar{s}_2/\bar{s}_1)^{1/(b-a)}$, $\theta = \bar{\delta} - [\bar{\delta}]$ *and*
$\bar{\theta} = (\bar{\delta}^{b-a} - (\bar{\delta} - \theta)^{b-a})/((\bar{\delta} + 1 - \theta)^{b-a} - (\bar{\delta} - \theta)^{b-a}) \in [0, 1)$.
$(0/0 = 0)$.
Then

$$\mathbf{P}(U_n) \geqslant \frac{\bar{\theta}\bar{s}_1^{b/(b-a)}}{\left(\bar{s}_2^{1/(b-a)} + (1-\bar{\theta})\bar{s}_1^{1/(b-a)}\right)^a} + \frac{(1-\bar{\theta})\bar{s}_1^{b/(b-a)}}{\left(\bar{s}_2^{1/(b-a)} - \bar{\theta}\bar{s}_1^{1/(b-a)}\right)^a}.$$

If $a = 1$ and $b = 2$, then Theorem 2 yields the Dawson–Sankoff inequality. For $a = 1$ and $b = 1.5$ the bound is better, but for $a, b \in \mathbb{N}$ moments $\bar{s}_i$ are sums of probabilities of intersections of the events.

4. Lower bounds for $\mathbf{P}(U_n)$.

Theorem (3)

*Let $\ell = 3$, $a > 0$, $\varrho > 0$. Put $f_{1i} = i^a$, $f_{2i} = i^{a+\varrho}$, $f_{ki} = i^{a+2\varrho}$ for all i. Put $\bar\delta_1 = n^\varrho \bar s_1 - \bar s_2$, $\bar\delta_2 = n^\varrho \bar s_2 - \bar s_3$,
$m = \min\{1 + [(\bar\delta_2/\bar\delta_1)^{1/\varrho}], n-1\}$ $(0/0 = 0)$. Then*

$$\mathbf{P}(U_n) \geqslant \frac{1}{n^a(m^\varrho - (m-1)^\varrho)}\Big(\frac{(m^\rho\bar\delta_1 - \bar\delta_2)(n^a - (m-1)^a)}{(m-1)^a(n^\varrho - (m-1)^\varrho)} - \frac{((m-1)^\rho\bar\delta_1 - \bar\delta_2)(n^a - m^a)}{m^a(n^\varrho - m^\varrho)}\Big) + \frac{\bar s_1}{n^a}.$$

For $a = 1$ and $\varrho = 1$, Theorem 3 is obtained by Kwerel (1975). In this case, moments $\bar s_i$ are sums of probabilities of intersections of the events.

We see two types of bounds. For moments of integer orders, they are good calculated and known. For non-integer orders, they may be stronger.

5. First jump of d.f. and Hölder's inequality.

Assume that $0 = x_0 < x_1 < x_2 < \cdots < x_n$. Let ξ be a random variable such that $p_i = \mathbf{P}(\xi = x_i)$ for $i = 0, 1, 2, \ldots, n$. Theorem 1 implies the following result.

Theorem (4)

Put $\bar{\delta} = (\bar{s}_2/\bar{s}_1)^{1/\varrho}$. *(We assume that* $0/0 = 0$.*) Let* m^* *be a natural number such that* $2 \leqslant m^* \leqslant n$ *and* $x_{m^*-1} \leqslant \bar{\delta} < x_{m^*}$. *For* $\bar{\delta} = 0$ *and* $\bar{\delta} = x_n$, *put* $m^* = n$.
Then

$$p_0 \leqslant 1 - \frac{\bar{s}_1}{x_{m^*}^{\varrho} - x_{m^*-1}^{\varrho}} \left(\frac{\bar{\delta}^{\varrho} - x_{m^*-1}^{\varrho}}{x_{m^*}^{a}} + \frac{x_{m^*}^{\varrho} - \bar{\delta}^{\varrho}}{x_{m^*-1}^{a}} \right) \leqslant 1 - \frac{\bar{s}_1^{(a+\varrho)/\varrho}}{\bar{s}_2^{a/\varrho}}.$$

p_0 is the probability of the first jump of d.f.

5. First jump of d.f. and Hölder's inequality.

Let $p > 1$ and $q > 1$ be numbers such that $\frac{1}{p} + \frac{1}{q} = 1$.

Theorem (5)

For every $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^n$ with $\boldsymbol{\alpha} \geqslant \mathbf{0}_n$, $\boldsymbol{\beta} \geqslant \mathbf{0}_n$ and $\|\boldsymbol{\alpha}\|_p = \|\boldsymbol{\beta}\|_q = 1$, there exist a natural number $\bar{n} \leqslant n$ and $\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}} \in \mathbb{R}^{\bar{n}}$ such that $\bar{\boldsymbol{\alpha}} \geqslant \mathbf{0}_{\bar{n}}$, $\bar{\boldsymbol{\beta}} \geqslant \mathbf{0}_{\bar{n}}$, $\|\bar{\boldsymbol{\alpha}}\|_p = \|\bar{\boldsymbol{\beta}}\|_q = 1$, $\bar{\boldsymbol{\alpha}}^T \bar{\boldsymbol{\beta}} = \boldsymbol{\alpha}^T \boldsymbol{\beta}$ and $x_1 < x_2 < \cdots < x_{\bar{n}}$ provided $\bar{n} > 1$, where $x_i = \bar{\alpha}_i \bar{\beta}_i^{-q/p}$ for $1 \leqslant i \leqslant \bar{n}$.
Assume that $\bar{n} > 1$. Put $\bar{s}_1 = \boldsymbol{\alpha}^T \boldsymbol{\beta}$.
Then there exists unique m^* with $2 \leqslant m^* \leqslant \bar{n}$ such that

$$1 \geqslant \bar{s}_1 \frac{x_{m^*}^p - x_{m^*-1}^p}{x_{m^*} x_{m^*-1}(x_{m^*}^{p-1} - x_{m^*-1}^{p-1})} - \frac{x_{m^*} - x_{m^*-1}}{x_{m^*} x_{m^*-1}(x_{m^*}^{p-1} - x_{m^*-1}^{p-1})} \geqslant \bar{s}_1^{p/(p-1)}. \quad (9)$$

$\bar{s}_1^{p/(p-1)} \leqslant 1$ is Hölder's inequality. The first inequality in (9) is

$$\sum_{i=1}^n \alpha_i \beta_i \leqslant \frac{x_{m^*} x_{m^*-1}(x_{m^*}^{p-1} - x_{m^*-1}^{p-1}) + x_{m^*} - x_{m^*-1}}{x_{m^*}^p - x_{m^*-1}^p}.$$

5. First jump of d.f. and Hölder's inequality.

In the proof of Theorem 5, the procedure of constructing of x_i is described.
For $p = 2$, there is an example in which there are two different bounds, better than that from Hölder's inequality.

6. Representations of $\mathbf{P}(U_n)$.

Numbers r_i in Theorem 1 may be chosen in any way to satisfy $\mathbf{P}(U_n) = \sum r_i$. Every representation of $\mathbf{P}(U_n)$ as a sum of non-negative numbers gives new bound.

Put $B_i = \{\xi_n = i\}$, $J_0 = \{0\}$ and
$J_r = \{j = (j_1, \ldots, j_r) : \; j_k \in \mathbb{N} \text{ and } 1 \leqslant j_k \leqslant n \text{ for all } 1 \leqslant k \leqslant r\}$
for $r \geqslant 1$. Then

$$\mathbb{I}_{U_n} = \sum_{i=1}^{n} \mathbb{I}_{B_i} = \sum_{i=1}^{n} \frac{\xi_n^r}{i^r} \mathbb{I}_{B_i} = \sum_{i=1}^{n} \sum_{j \in J_r} \frac{1}{i^r} \mathbb{I}_{B_i A_{j_1} \ldots A_{j_r}}.$$

Lemma

Fix integer r with $0 \leqslant r \leqslant n$. For $r \geqslant 1$, put $p_{i,j} = \mathbf{P}(B_i A_{j_1} \ldots A_{j_r})$, $j \in J_r$. For $r = 0$, put $p_{i,j} = \mathbf{P}(B_i)$ for all $j \in J_0$.
Then

$$\mathbf{P}(U_n) = \sum_{i=1}^{n} \sum_{j \in J_r} \frac{1}{i^r} p_{i,j} = \sum_{j \in J_r} R_j.$$

Theorem 1 yield bounds for every R_j.

7. Further lower bounds for $\mathbf{P}(U_n)$.

Put $r = 1$. Denote $r_{ik} = p_{i,j}/i$ and

$$\bar{s}_k(j) = \sum_{i=1}^{n} i^{a+(k-1)\varrho} r_{ij}, \quad 1 \leqslant k \leqslant \ell,\ j = 1, 2, \ldots, n.$$

Theorem (7)

Put $\ell = 2$, $\bar{\delta}_j = (\bar{s}_2(j)/\bar{s}_1(j))^{1/\varrho}$, $\theta_j = \bar{\delta}_j - [\bar{\delta}_j]$ *and* $\bar{\theta}_j = (\bar{\delta}_j^{\varrho} - (\bar{\delta}_j - \theta_j)^{\varrho})/((\bar{\delta}_j + 1 - \theta_j)^{\varrho} - (\bar{\delta}_j - \theta_j)^{\varrho}) \in [0, 1)$, *where* $j = 1, 2, \ldots, n$.
Then

$$\mathbf{P}(U_n) \geqslant \sum_{j=1}^{n} \left\{ \frac{\bar{\theta}_j \bar{s}_1^{(a+\varrho)/\varrho}(j)}{\left(\bar{s}_2^{1/\varrho}(j) + (1-\theta_j)\bar{s}_1^{1/\varrho}(j)\right)^a} + \frac{(1-\bar{\theta}_j)\bar{s}_1^{(a+\varrho)/\varrho}(j)}{\left(\bar{s}_2^{1/\varrho}(j) - \theta_j \bar{s}_1^{1/\varrho}(j)\right)^a} \right\}. \quad (10)$$

For $a = \varrho = 1$, Theorem 7 implies a result of Kuai, Alajaji and Takahara (2000) which generalizes that of de Caen (1997).

7. Further lower bounds for $\mathbf{P}(U_n)$.

For $\ell = 3$, a formula is large. No earlier results. The same is for upper bounds.
The case $r \geqslant 2$ did not considered earlier.

8. Lower bounds for $\mathbf{P}(U_n|\mathcal{A})$.

Let $\mathcal{A}$ be a σ-field of events. The above method works for $\mathbf{P}(U_n|\mathcal{A})$ as well. For example, we have an analogue of previous theorem.

Theorem (8)

For $j \in J_m$, $m \geqslant 0$, define random variables

$$\bar{\delta}^{\mathcal{A}}(j) = \left(\frac{\bar{s}_2^{\mathcal{A}}(j)}{\bar{s}_1^{\mathcal{A}}(j)}\right)^{1/\varrho}, \quad \theta^{\mathcal{A}}(j) = \bar{\delta}^{\mathcal{A}}(j) - [\bar{\delta}^{\mathcal{A}}(j)],$$

$$\bar{\theta}^{\mathcal{A}}(j) = \frac{(\bar{\delta}^{\mathcal{A}}(j))^{\varrho} - (\bar{\delta}^{\mathcal{A}}(j) - \theta^{\mathcal{A}}(j))^{\rho}}{(\bar{\delta}^{\mathcal{A}}(j) + 1 - \theta^{\mathcal{A}}(j))^{\varrho} - (\bar{\delta}^{\mathcal{A}}(j) - \theta^{\mathcal{A}}(j))^{\rho}},$$

Note that $\bar{\theta}^{\mathcal{A}}(j) \in [0,1)$ a.s. for all $j \in J_m$. Then w.p. 1

$$\mathbf{P}(U_n|\mathcal{A}) \geqslant \sum_{j \in J_m} \Big(\frac{\bar{\theta}^{\mathcal{A}}(j)(\bar{s}_1^{\mathcal{A}}(j))^{(a+\varrho)/\varrho}}{\left((\bar{s}_2^{\mathcal{A}}(j))^{1/\varrho} + (1 - \theta^{\mathcal{A}}(j))(\bar{s}_1^{\mathcal{A}}(j))^{1/\varrho}\right)^a} + \frac{(1 - \bar{\theta}^{\mathcal{A}}(j))(\bar{s}_1^{\mathcal{A}}(j))^{(a+\varrho)/\varrho}}{\left((\bar{s}_2^{\mathcal{A}}(j))^{1/\varrho} - \theta^{\mathcal{A}}(j)(\bar{s}_1^{\mathcal{A}}(j))^{1/\varrho}\right)^a} \Big).$$

8. Lower bounds for $\mathbf{P}(U_n|\mathcal{A})$.

Taking the expectation from both sides of such bounds, we get new bound for unconditional $\mathbf{P}(U_n)$.

All previous techniques works for probabilities that at least p from n events occur. (Conditional and ususal, upper and lower, with similar representations).

Probability may be replaced by a measure or a measure with sign.

9. The Borel–Cantelli lemma.

Borel (1909).
Erdős and Rényi (1959):

Let $\{A_n\}$ be a sequence of events such that $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$. Put

$$L = \liminf_{n\to\infty} \frac{\sum_{i,j=1}^{n} \mathbf{P}(A_i A_j)}{\left(\sum_{i=1}^{n} \mathbf{P}(A_i)\right)^2}.$$

If $L = 1$, then $\mathbf{P}(A_n\ i.o.) = \mathbf{P}(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = 1$.
Kochen and Stone (1964) and Spitzer (1964) obtained the following generalization of this result: $\mathbf{P}(A_n\ i.o.) \geqslant 1/L$.

9. The Borel–Cantelli lemma.

Further generalizations: Kounias (1968), Móri and Székely (1983), Martikainen and Petrov (1990), Petrov (2002, 2004), Andel and Dupas (1989), Feng, Li and Shen (2009), Frolov (2012, 2015).

For $m \leqslant n$, put $U_{mn} = \bigcup\limits_{k=m}^{n} A_k$. Since

$$\mathbf{P}\left(A_n \text{ i.o.}\right) = \mathbf{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = \lim_{m\to\infty} \lim_{n\to\infty} \mathbf{P}\left(U_{mn}\right),$$

every new upper or lower bound allows us to derive new variant of first or second part of the Borel–Cantelli Lemma.

9. The Borel–Cantelli lemma.

Theorem (9)

Let $\{A_i\}$ *be a sequence of events such that* $\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \infty$*. Put*

$$s_1(n) = \sum_{i=1}^{n} \mathbf{P}(A_i),\ s_2(n) = 2\sum_{1\leqslant i<j\leqslant n} \mathbf{P}(A_iA_j),\ s_3(n) = 6\sum_{1\leqslant i<j<k\leqslant n} \mathbf{P}(A_iA_jA_k),$$

$\delta_1(n) = (n-1)s_1(n) - s_2(n)$, $\delta_2(n) = (n-2)s_2(n) - s_3(n)$.
Assume that $\delta_1(n)/n \to \infty$ *and* $s_2(n) = o(\delta_1(n) + \delta_2(n))$ *as* $n \to \infty$.
Then

$$\mathbf{P}(A_i \ i.o.) \geqslant \limsup_{n\to\infty} \left(\frac{(\delta_1(n))^2}{n(\delta_1(n)+\delta_2(n))} + \frac{s_1(n)}{n} \right). \tag{11}$$

9. The Borel–Cantelli lemma.

Theorem 9 is from Frolov (2012). The bound may be better than those from previous generalizations. The proof uses Theorem 3. From Theorem 7,

Theorem (10)

Denote $\xi_n = I_{A_1} + I_{A_2} + \cdots + I_{A_n}$ and $\eta_n = n - \xi_n$ for all natural n. Assume that

$$\frac{1}{n} \sum_{k=1}^{n} \frac{\mathbf{E}\eta_n I_{A_k}}{\mathbf{E}\eta_n \xi_n I_{A_k}} \to 0 \quad as \quad n \to \infty.$$

Then

$$\mathbf{P}\left(A_n \ i.o.\right) \geqslant \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left\{ \mathbf{P}\left(A_k\right) + \frac{(\mathbf{E}\eta_n I_{A_k})^2}{\mathbf{E}\eta_n \xi_n I_{A_k}} \right\}.$$

Here $\mathbf{E}\eta_n I_{A_k} = \sum_{i=1}^{n} \mathbf{P}\left(\overline{A_i} A_k\right)$, $\quad \mathbf{E}\eta_n \xi_n I_{A_k} = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{P}\left(\overline{A_i} A_j A_k\right).$

9. The Borel–Cantelli lemma.

Conditional variants are obtained as well. This is a conditional variant of a result in Petrov (2002).

Theorem (11)

Put $D = \left\{ \omega : \sum_{k=1}^{\infty} \mathbf{P}(A_k|\mathcal{A}) = \infty \right\}$. *Assume that* $\mathbf{P}(A_k A_j|\mathcal{A}) \leqslant \zeta_n \mathbf{P}(A_k|\mathcal{A})\,\mathbf{P}(A_j|\mathcal{A})$ *for almost all* $\omega \in D$ *and all* $k \neq j$, $1 \leqslant k, j \leqslant n$, *where* $\{\zeta_n\}$ *is a sequence of random variables such that* $\zeta_n \geqslant 1$ *for almost all* $\omega \in D$ *and all* n.
Then

$$\mathbf{P}(\limsup A_n|\mathcal{A}) \geqslant \limsup_{n\to\infty} \frac{1}{\zeta_n}$$

for almost all $\omega \in D$.

For trivial $\mathcal{A}$ and degenerated ζ_n Theorem 11 turns to the result in Petrov (2002).

The end.

Thank you for your attention.