

МАТЕМАТИКА ВЫБОРОВ

Гасников А.В.,

доцент кафедры МОУ ФУПМ МФТИ

Черноусова Е.О.,

аспирант кафедры ИС ВЦ РАН

ПЛАН ВЫСТУПЛЕНИЯ

- 1. Этап предвыборной кампании, влияние СМИ на общественное мнение**
- 2. Опросы общественного мнения**
- 3. Возможные особенности результатов голосования**

ЧЕСТНАЯ ЖЕРЕБЬЕВКА

Пусть в некоторой телевизионной программе (согласно предвыборной кампании) n партиям разрешается в течение фиксированного времени выступить с агитационной речью. Выбор того, в каком порядке будут выступать представители n партий, решается жеребьевкой.

Предложите процедуру честной жеребьевки.

НЕДОВЕРЧИВЫЕ ИГРОКИ

Два игрока (А и В) хотят сыграть в орлянку. У каждого из них есть монета. Но они не доверяют друг, подозревая, что у соперника несимметричная монета.

Как быть?

Решение:

Кинуть обе монеты сразу: если обе монеты выпали орлом или обе решкой, то выиграл игрок А, если по-разному – то игрок В.

Если обе монеты независимы и **хотя бы одна** из них симметрична, то вероятность выигрыша в такой игре равна $\frac{1}{2}$.

Пусть игрок А знает, что его монета симметрична ($p_A = \frac{1}{2}$), но не знает p_B .

Тогда вероятность выигрыша для игрока А равна

$$p_A p_B + (1 - p_A)(1 - p_B) = \frac{1}{2} p_B + \frac{1}{2} (1 - p_B) = \frac{1}{2}$$

ЧЕСТНАЯ ЖЕРЕБЬЕВКА

Перестановкой на n элементах называется биекция множества $\{1, 2, \dots, n\}$ на себя: $\sigma_n = (i_1, i_2, \dots, i_n) \Leftrightarrow \sigma_n(j) = i_j$.

Число различных перестановок равно $n!$

Случайная перестановка. Вероятность реализации конкретной перестановки равна $1/n!$ (равномерная мера на множестве всех перестановок длины n).

Пусть есть некоторый набор перестановок $\sigma_n^1, \sigma_n^2, \dots, \sigma_n^k$. Известно, что хотя бы одна из них случайная. Тогда перестановка, получающаяся в результате последовательного перемножения каждой из них (не важно, в каком порядке), будет случайной.

ЧЕСТНАЯ ЖЕРЕБЬЕВКА

Решение:

Представитель каждой партии приносит на телевизионную программу некоторую перестановку; эти конверты одновременно вскрываются в присутствии всех участников и затем последовательно (скажем, в алфавитном порядке названий партий) выполняются указанные в них перестановки.

Процедура гарантирует, что если **хотя бы одна** из партий соблюдает правила (ее перестановка случайна), то результат жеребьевки будет **честным**.

КАЖДЫЙ ПО ОТДЕЛЬНОСТИ – ПРОТИВ, А ВСЕ ВМЕСТЕ – ЗА (ИЛИ НАОБОРОТ)

Осторожная привычка людей(избирателей) вести себя «как все».

Предлагается для обсуждения простая версия подобного механизма.

Считается, что индивидuum (i), принимая решение, руководствуется:

- 1) личным (априорным) отношением к данному вопросу - a_i ;**
- 2) отношением к этому вопросу окружающих его субъектов (коллектива).**

Финальное, апостериорное отношение, сформировавшееся после общения с коллективом, определяется числом p_i зависит от пп. 1, 2.

μ_i - мера независимости индивидуума - вероятность, что в конкретной ситуации индивидuum ведет себя как независимый.

Будем считать, что влияние каждого j -го члена коллектива на данного i -го индивидуума не зависит от влияния любого другого из остальных членов и определяется числом λ_{ij} - вероятностью, что i -ый поступит также как j -ый. $\lambda_{ii} = 0$

По формуле полной вероятности:

$$p_i = a\mu_i + (1 - \mu_i) \sum_{j=1}^N \lambda_{ij} p_j$$

Или в матричной форме:

$$p = aM + (I_N - M)\Lambda p$$

ТЕОРЕМА

Решение p всегда существует, неотрицательно и каждое p_i не превосходит единицу. Решение p единственно, если не все $\mu_i = 0$

СТАДО БЕЗ ВОЖАКА

Все $\mu_i = 0$ (т.е. все индивидуумы абсолютно зависимы) $(I_N - \Lambda)p = 0$

Так как $\det(I_N - \Lambda) = 0$, решение существует и неединственное.

Несмотря на неопределенность состояния коллектива, индивидуумы ведут себя как единое целое: все p_i равны между собой, т.е. индивидуумы подражают друг другу. Коллектив легко выводится из такого «безразличного» состояния любым провоцирующим действием (малым возмущением одного из параметров μ_i). Пусть $\mu_j > 0$, тогда по теореме решение системы $p = aM + (I_N - M)\Lambda p$ единственно: у всех индивидуумов $p_i = a_j$ (т.е. поведение j -го индивидуума копируют все).

Дурная овца все стадо портит (русская поговорка)

ТОК-ШОУ

Пусть $p_i = a_i \mu_i + (1 - \mu_i) \sum_{j=1}^N \frac{1 - \delta_{ij}}{N - 1} p_j$, $i = 1, \dots, N$.

$\sum_{j=1}^N \frac{1 - \delta_{ij}}{N - 1} p_j$ - математическое ожидание доли индивидуумов – зрителей в студии - (помимо i -го), перешедших в данное состояние.

Решение такой системы (при больших значениях N):

$$p_i = a_i \mu_i + N \frac{1 - \mu_i}{N - \mu_i} \frac{\sum_{j=1}^N a_j \mu_j}{\sum_{j=1}^N \mu_j}, \quad \frac{M}{N} = \frac{\sum_{j=1}^N a_j \mu_j}{\sum_{j=1}^N \mu_j}$$

$M = \sum_{j=1}^N p_j$ - математическое ожидание числа индивидуумов, перешедших в данное состояние.

ТОК-ШОУ

Пусть выступают два оппонента: у первого - $a_1 = 1, \mu_1 > 0$; у второго - $a_2 = 0, \mu_2 > 0$.

У зрителей в студии все $\mu_i = 0$, а разброс мнений достаточно широк, т.е. a_i - разные.

В результате зрители разделится на две части (в независимости от своих априорных убеждений) в отношении $\mu_1 : \mu_2$.

МОДЕЛЬ ВЛИЯНИЯ ТЕЛЕВИЗОРА НА МНЕНИЕ ЛЮДЕЙ

Элементарная общественная ячейка (коллектив взаимных влияний) - семья

Модель:
$$p_i = a_i \mu_i + (1 - \mu_i) \sum_{j=0}^N \frac{1 - \delta_{ij}}{N} p_j$$

Параметры:

телевизор: $\mu_0 = p_0 = a_0 = 1$

остальные члены семьи: $p_i = \bar{p}, \mu_i = \bar{\mu}, a_i = \bar{a}$

$$\bar{p} = \bar{a} \bar{\mu} + \frac{(1 - \bar{\mu})(1 + (N - 1)\bar{p})}{N}$$

Решение:
$$\bar{p} = \frac{M}{N} = \frac{N\bar{a}\bar{\mu} + (1 - \bar{\mu})}{1 + (N - 1)\bar{\mu}}$$

при $N = 4$ получаем
$$\bar{p} = \frac{M}{4} = \frac{4\bar{a}\bar{\mu} + (1 - \bar{\mu})}{1 + 3\bar{\mu}}$$

МОДЕЛЬ ВЛИЯНИЯ ТЕЛЕВИЗОРА НА МНЕНИЕ ЛЮДЕЙ

Пусть было три этапа предвыборной кампании.

1) Положим $\bar{\mu} = \frac{1}{2}$, $\bar{p}^{(0)} = 0$, $\bar{a}^{(0)} = 0$, тогда $\bar{p}^{(1)} = 0.2$;

2) Положим $\bar{a}^{(1)} = \bar{p}^{(1)}$, тогда $\bar{p}^{(2)} = 0.36$;

3) Положим $\bar{a}^{(2)} = \bar{p}^{(2)}$, тогда $\bar{p}^{(3)} = 0.48$;

СОЦИОЛОГИЧЕСКИЙ ОПРОС ИЛИ EXIT POLL

Пусть в некоторой области прошел второй тур выборов. Выбор был между двумя кандидатами (А и В). Сколько человек надо опросить на выходе с избирательных участков, чтобы определить процент проголосовавших за кандидата А с точностью 1% и с (доверительной) вероятностью не менее 0.95.

Решение:

Пусть число жителей города, участвующих в голосовании, равно N . Из них M человек проголосовало за кандидата А. Обозначим за n – общее число опрошенных человек (в ходе exit-polla), среди них m – число отдавших свой голос за кандидата А.

$$P\left\{\left|\frac{m}{n} - \frac{M}{N}\right| > 0.01\right\} < 1 - 0.95$$

СОЦИОЛОГИЧЕСКИЙ ОПРОС ИЛИ EXIT POLL

При фиксированном n случайная величина m (число отдавших свой голос за кандидата А среди опрошенных) имеет

гипергеометрическое распределение $\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}$.

Формула Стирлинга: $N! = \sqrt{2\pi N} (Ne^{-1})^N (1 + o(N))$

Если $n = o(\sqrt{N})$, то $C_N^n \approx \frac{N^n}{n!}$.

Тогда гипергеометрическое распределение $G(M, N, n)$ аппроксимируется

биномиальным $Bi\left(\frac{M}{N}, n\right) \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \approx C_n^m \left(\frac{M}{N}\right)^m \left(1 - \frac{M}{N}\right)^{n-m}$.

СОЦИОЛОГИЧЕСКИЙ ОПРОС ИЛИ EXIT POLL

Граница Чебышева: $P\left\{\left|\xi - E\xi\right| > \varepsilon\right\} \leq \frac{D\xi}{\varepsilon^2}$

$$P\left\{\left|\frac{\sum_{i=1}^n X_i}{n} - p\right| < \varepsilon\right\} \leq \frac{DX}{\varepsilon^2 n}$$

$$P\left\{\left|\frac{m}{n} - \frac{M}{N}\right| > 0.01\right\} \leq \frac{\frac{M}{N}\left(1 - \frac{M}{N}\right)}{n(0.01)^2} \leq \frac{0.25}{n(0.01)^2}$$

$$\frac{0.25}{n(0.01)^2} \leq 1 - 0.95 \Rightarrow n \geq 50000$$

СОЦИОЛОГИЧЕСКИЙ ОПРОС ИЛИ EXIT POLL

Граница Чернова:

$$P \left\{ \sum_{i=1}^n X_i > (p + \varepsilon)n \right\} \leq \exp \left\{ -2\varepsilon^2 n \right\}$$

$$P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - p \right| > \varepsilon \right\} \leq 2 \exp \left\{ -2\varepsilon^2 n \right\}$$

$$P \left\{ \left| \frac{m}{n} - \frac{M}{N} \right| > 0.01 \right\} \leq 2 \exp \left\{ -2(0.01)^2 n \right\}$$

$$2 \exp \left\{ -2(0.01)^2 n \right\} < 1 - 0.95 \Rightarrow n \geq 10000$$

СОЦИОЛОГИЧЕСКИЙ ОПРОС ИЛИ EXIT POLL

Центральная предельная теорема

$$P \left\{ \left| \frac{\sum_{i=1}^n X_i - np}{\sqrt{npq}} \right| > \varepsilon \right\} \approx 2\Phi(\varepsilon) - 1, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

$$P \left\{ \frac{\left| m - n \frac{M}{N} \right|}{\sqrt{n \frac{M}{N} \left(1 - \frac{M}{N} \right)}} > \frac{\sqrt{n}}{\sqrt{\frac{M}{N} \left(1 - \frac{M}{N} \right)}} 0.01 \right\} \approx 2\Phi \left(\frac{\sqrt{n}}{\sqrt{\frac{M}{N} \left(1 - \frac{M}{N} \right)}} 0.01 \right) - 1$$

$$P \left\{ \frac{\left| \sum_{i=1}^n X_i - np \right|}{\sqrt{npq}} > \varepsilon \right\} \approx 2(1 - \Phi(\varepsilon)) \Rightarrow \varepsilon \approx 1.96 \Rightarrow \frac{\sqrt{n}}{\sqrt{\frac{M}{N} \left(1 - \frac{M}{N} \right)}} 0.01 \approx 1.96 \Rightarrow n \approx 9600$$

СОЦИОЛОГИЧЕСКИЙ ОПРОС ИЛИ EXIT POLL

Заметим, что результат задачи не зависит от общего числа избирателей, кроме как от условия $n = o(\sqrt{N})$.

Если хотим сделать более грубую оценку доли проголосовавших за кандидата А, например, с точностью всего 5%, то достаточно опросить всего ~390 человек!!!

$$\frac{\sqrt{n}}{\sqrt{\frac{M}{N} \left(1 - \frac{M}{N}\right)}} 0.05 \approx 1.96 \Rightarrow n \approx 390$$

Но возникает вопрос **репрезентативности** выборки из опрошенных избирателей.

АНАЛИЗ СТАТИСТИЧЕСКИХ РЕЗУЛЬТАТОВ И СПЕЦИФИКА, В Т.Ч. НЕОДНОРОДНОСТЬ, ДАННЫХ

Для условного соответствия тенденциям в электоральном поведении населения можно выделить:

- А) «городские» участки;**
- Б) «сельские» участки;**
- В) «особые» участки (больницы, тюрьмы, дома престарелых, военные городки);**

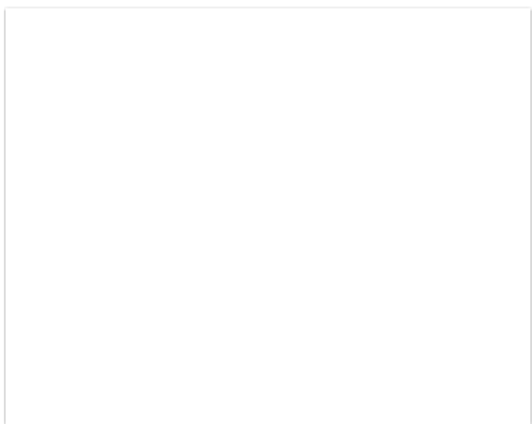
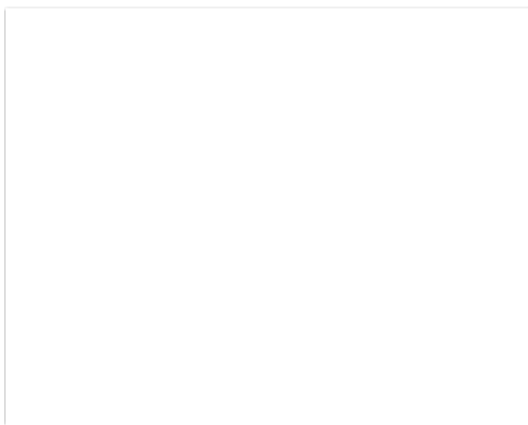
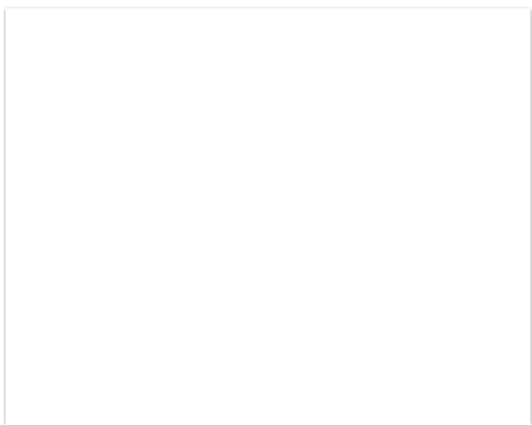
ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

$\{\xi_n\}$ - последовательность независимых **одинаково** распределенных с.в.

$$D\xi_k = \sigma^2 \neq 0$$

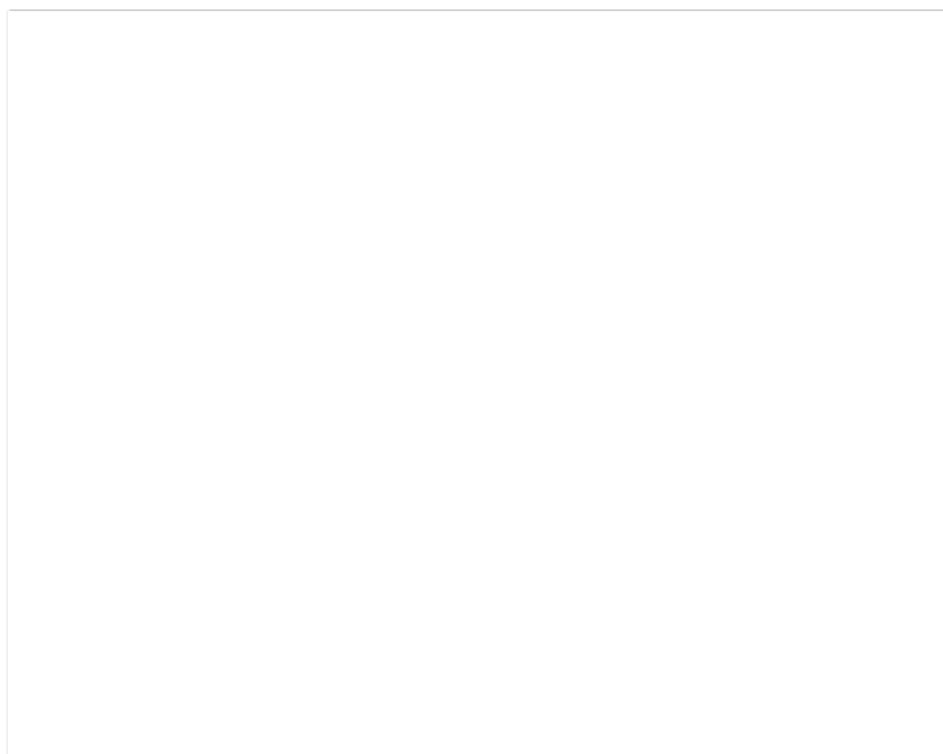
$$\frac{\sum_{k=1}^n (\xi_k - E\xi_k)}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1) \quad \Leftrightarrow \quad \frac{1}{n} \sum_{k=1}^n \xi_k \xrightarrow{d} N\left(E\xi_k, \sigma^2/n\right)$$

Рассмотрим $\{\xi_n\}$, $\{\eta_n\}$, где $\xi_n : Be(0.35)$, $\eta_n : Be(0.5)$



ПОЯВЛЕНИЕ «ПИКОВ» ПРИ МАЛЫХ РАЗМЕРАХ ИЗБИРАТЕЛЬНОГО УЧАСТКА

Повторяемость определенных долей явки/голосов избирателей:



СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. П.С.Краснощеков, А.А.Петров «Принципы построения моделей» Изд. 2-е – М.: ФАЗИС: ВЦ РАН 2000.
2. А.Шень «Вероятность: примеры и задачи» - М.: МЦНМО, 2007
3. Материалы из «Живого журнала»

**СПАСИБО
ЗА ВНИМАНИЕ!**


```
numb_s_uik=10000;  
size_s_uik=unidrnd(500, numb_s_uik, 1);  
javka_temp=unifrnd(0.6, 0.9, numb_s_uik, 1);  
javka_s_uik=binornd(size_s_uik, javka_temp);  
itog_javka=javka_s_uik./size_s_uik;  
golos_temp=unifrnd(0.6, 0.9, numb_s_uik, 1);  
golos_s_uik=binornd(javka_s_uik, golos_temp);  
I_s=hygernd(size_s_uik,golos_s_uik,javka_s_uik);  
hist(I_s./javka_s_uik,1000)
```