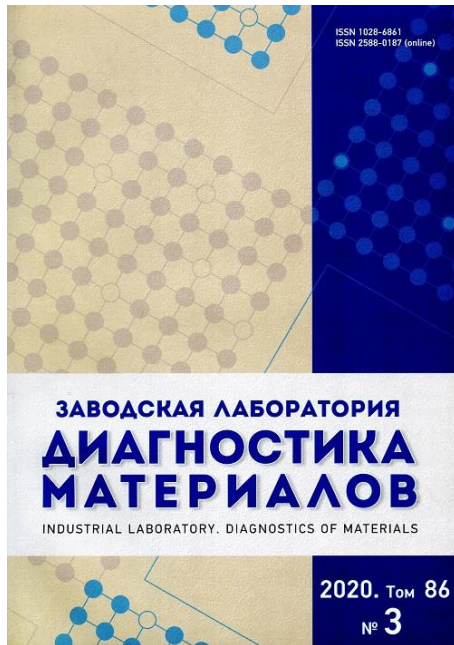


Александр Орлов. О методах проверки однородности двух независимых выборок

В последнее время я [изучаю](#) решения менеджеров в Fantasy Premier League (FPL). В частности, есть ли отличия в поведении лучших менеджеров, отобранных по результатам предыдущих сезонов, и случайной выборкой. Важный аспект такого исследования – вынесение суждения, однородны ли выборки. Можно ли объяснить различия случайностью, или они закономерны? Ранее я рассказал о [критерии Колмогорова](#) и трудностях его использования. Настоящая статья Александра Ивановича Орлова дает современный обзор методов проверки однородности двух независимых выборок. По ходу изложения я применяю критерии к задаче FPL. Этот текст набран с отступом.

Александр Иванович Орлов. О методах проверки однородности двух независимых выборок. – Заводская лаборатория. Диагностика материалов. 2020. Том 86. № 3. Стр. 67–76.



[Оригинал](#) статьи (для загрузки требуется авторизация).

Базовая вероятностно-статистическая модель

Постановка задачи в математико-статистических терминах: имеются две выборки x_1, x_2, \dots, x_m и y_1, y_2, \dots, y_n , требуется проверить их однородность. Понятием, противоположным *однородности*, является *различие* (или *наличие эффекта*). Можно переформулировать задачу: требуется проверить, есть ли различие между выборками.

При проверке однородности двух выборок общепринята модель, в которой x_1, x_2, \dots, x_m рассматриваются как результаты m независимых наблюдений некоторой случайной величины X с функцией распределения $F(x)$, неизвестной статистике, а y_1, y_2, \dots, y_n – как результаты n независимых наблюдений, вообще говоря, другой случайной величины Y с функцией распределения $G(x)$, также неизвестной статистике. Кроме того, предполагается, что наблюдения в одной выборке не зависят от наблюдений в другой, поэтому выборки и называют независимыми.

В [Fantasy Premier League](#) сезона 2022/23 участвует более 10М аккаунтов. Я сформировал две выборки. *Элита* – ТОП-10К по рейтингу за пять предыдущих сезонов. *Поляна* – 10К, полученных с помощью генератора случайных чисел из первых 5М аккаунтов.

Основные постановки задачи проверки однородности двух независимых выборок

Однородность достигается, если обе выборки взяты из одной генеральной совокупности, т. е. справедлива нулевая гипотеза $H_0: F(x) = G(x)$ при всех x . Отсутствие однородности означает, что верна альтернативная гипотеза $H_1: F(x_0) \neq G(x_0)$ хотя бы при одном значении аргумента x_0 .

В некоторых случаях целесообразно проверять совпадение не функций распределения, а лишь некоторых характеристик случайных величин X и Y – математических ожиданий, медиан, дисперсий. Например, однородность математических ожиданий означает, что справедлива гипотеза $H'_0: M(X) = M(Y)$, где $M(X)$ и $M(Y)$ – математические ожидания случайных величин X и Y .

Проверка однородности характеристик

Для проверки однородности математических ожиданий традиционно используют двухвыборочный критерий Стьюдента. К настоящему времени этот метод устарел, но по традиции встречается в учебной литературе. Сначала вычисляют выборочные средние арифметические

$$\bar{x} = \frac{1}{m} \sum_{1 \leq i \leq m} x_i, \quad y = \frac{1}{n} \sum_{1 \leq i \leq n} y_i$$

затем — выборочные дисперсии

$$s_x^2 = \frac{1}{m-1} \sum_{1 \leq i \leq m} (x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{n-1} \sum_{1 \leq i \leq n} (y_i - \bar{y})^2$$

и статистику Стьюдента

$$(1) \quad t = \frac{\bar{x} - \bar{y}}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2}} \sqrt{\frac{mn(m+n-2)}{m+n}}$$

на основе которой принимают решение.

По заданному уровню значимости α и числу степеней свободы $(m+n-2)$ из таблиц распределения Стьюдента находят критическое значение $t_{кр}$. Если $|t| > t_{кр}$, то гипотезу однородности (отсутствия различия) отклоняют, если же $|t| < t_{кр}$, то — принимают.

Согласно теории, должны быть выполнены два условия применимости критерия Стьюдента, основанного на использовании статистики t , заданной формулой (1): результаты наблюдений имеют нормальные распределения и дисперсии результатов наблюдений в первой и второй выборках совпадают.

Если эти условия выполнены, то статистика t при справедливости H_0 имеет распределение Стьюдента с $(m+n-2)$ степенями свободы. Только в этом случае описанный выше традиционный метод обоснован. Если хотя бы одно из условий не выполнено, то нет оснований считать, что статистика t имеет распределение Стьюдента.

Классические условия применимости критерия Стьюдента в подавляющем большинстве технических, экономических, медицинских и иных задач не выполнены. Тем не менее при больших и примерно равных объемах выборок его можно применять. При конечных объемах выборок традиционный метод носит неустраимо приближенный характер.

Я подсчитал очки, набранные менеджерами в двух выборках по итогам девятнадцати игровых недель, среднее \bar{x} , дисперсию s^2 и критерий Стьюдента t . Поскольку $m = n$, (1) можно переписать в виде:

$$(2) \quad t = \frac{\bar{x} - \bar{y}}{\sqrt{s_x^2 + s_y^2}} \cdot \sqrt{n}$$

	A	B	C	D	E	F	G	H	I	J	K	L
1	элита			поляна					элита	поляна		
2	аккаунт	очки		аккаунт	очки		среднее	\bar{x}	1108	1070	=СРЗНАЧ(поляна[очки])	
3	16 409	809		4 818 840	303		дисперсия	s^2	1735	3524	=ДИСП.В(поляна[очки])	
4	11 544	814		30 063	385		размер выборки	n	10000	10000	=СЧЁТЗ(поляна[очки])	
5	423 487	823		1 386 823	589		статистика Стьюдента	t		51	=(I2-J2)*КОРЕНЬ(I4)/КОРЕНЬ(I3+J3)	
6	7 273	825		460 636	779		число степеней свободы	df		19998	=I4+J4-2	
7	150 155	850		4 340 012	797		уровень значимости	α		0,05		
8	52 233	947		928 358	798			$t_{кр}$		1,9601	=СТЮДЕНТ.ОБР.2Х(I7;J6)	
9	4 726 349	949		3 953 627	800							
10	1 394	955		4 201 358	805							
11	11 802	956		3 982 646	808							
12	22 119	958		12 032	828							
13	61 368	959		435 608	828							
14	3 084 777	960		576 011	834							
15	1 622 232	963		751 357	834							

$$(2) \quad t = \frac{\bar{x} - \bar{y}}{\sqrt{s_x^2 + s_y^2}} \cdot \sqrt{n}$$

Рис. 1. Двухвыборочный критерий Стьюдента (см. также приложенный файл Excel)

Статистика Стьюдента $t = 51$ существенно больше $t_{кр}=1,96$. Это означает, что гипотезу однородности следует отклонить. Результаты элиты и поляны различаются.

Критерий Крамера – Уэлча равенства математических ожиданий

Вместо критерия Стьюдента для проверки H_0 целесообразно использовать критерий Крамера – Уэлча, основанный на статистике

$$(6) T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{m} + \frac{s_y^2}{n}}} = \frac{\sqrt{mn}(\bar{x} - \bar{y})}{\sqrt{ns_x^2 + ms_y^2}}$$

Критерий Крамера – Уэлча имеет прозрачный смысл – разность выборочных средних арифметических для двух выборок делится на естественную оценку среднего квадратического отклонения этой разности. Естественность указанной оценки состоит в том, что неизвестные статистику дисперсии заменены их выборочными оценками. Из [Центральной предельной теоремы](#) и из [теорем о наследовании сходимости](#) следует, что при росте объемов выборок распределение статистики T Крамера – Уэлча сходится к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1: $\Phi(x) = N(x; 0, 1)$.

При $m = n$, как следует из формул (1) и (6), $t = T$. Если $M(X) \neq M(Y)$, то при больших объемах выборок

$$(7) P(T \leq X) \approx \Phi(x - c_{mn})$$

где

$$(8) c_{mn} = \frac{\sqrt{mn}[M(X) - M(Y)]}{\sqrt{nD(X) + mD(Y)}}$$

Правило принятия решения для критерия Крамера – Уэлча: если $|T| \leq \Phi^{-1}(1 - \alpha/2)$, то гипотеза однородности математических ожиданий принимается на уровне значимости α ; если $|T| > \Phi^{-1}(1 - \alpha/2)$, то гипотеза однородности математических ожиданий отклоняется на уровне значимости α .

Для $\alpha = 0,05$ значение модуля статистики T Крамера – Уэлча надо сравнивать с граничным значением $\Phi^{-1}(1 - \alpha/2) = 1,96$.

Таким образом при анализе организационно-экономических данных более обосновано применение критерия Крамера – Уэлча, чем критерия Стьюдента. Дополнительное преимущество критерия Крамера – Уэлча по сравнению с критерием Стьюдента – не требуется равенства дисперсий $D(X) = D(Y)$. Распределение статистики T не является распределением Стьюдента, однако и распределение статистики t , как показано выше, не является таковым в реальных ситуациях.

Поскольку размер двух выборок FPL одинаков, $m = n$, как было сказано выше $t = T$, и статистика Крамера – Уэлча ничего не добавляет к критерию Стьюдента.

Непараметрические методы проверки однородности

В большинстве задач представляет интерес не проверка равенства математических ожиданий или иных характеристик распределения, а обнаружение различия генеральных совокупностей, из которых извлечены выборки, т. е. проверка гипотезы H_0 . И здесь статистики t Стьюдента и T Крамера – Уэлча не годятся. Априорное предположение о принадлежности функций распределения $F(x)$ и $G(x)$ к какому-либо определенному параметрическому семейству обычно нельзя достаточно надежно обосновать. Поэтому для проверки H_0 следует использовать методы, пригодные при любом виде $F(x)$ и $G(x)$, т. е. непараметрические методы. *Непараметрический* означает, что нет необходимости предполагать вид функции распределения результатов наблюдений.

Двухвыборочный критерий Вилкоксона

Этот критерий (в литературе его называют также критерием Манна – Уитни) предназначен для проверки гипотезы

$$(9) H_{0m}: P(X < Y) = 1/2$$

где X — случайная величина, распределенная как элементы первой выборки, а Y — случайная величина, распределенная как элементы второй выборки. Альтернативой является отрицание H_{0m} :

вероятность $P(X < Y)$ отлична от 0,5. Это — непараметрическая гипотеза. Но из нее не следует, что функции распределения двух выборок совпадают.

Критерий Вилкоксона – один из самых известных инструментов непараметрической статистики. При альтернативной гипотезе, когда функции распределения выборок $F(x)$ и $G(x)$ не совпадают, распределение статистики Вилкоксона зависит от величины $a = P(X < Y)$. Если a отличается от 1/2, то мощность критерия Вилкоксона стремится к 1 и он отличает нулевую гипотезу $F \equiv G$ от альтернативной.

Можно переформулировать нулевую гипотезу в терминах равенства медиан двух выборок:

$$(B1) H_0: M_1 = M_2$$

Тогда альтернативная двусторонняя гипотеза, просто отрицает равенство медиан

$$(B2) H_1: M_1 \neq M_2$$

При равенстве размера двух выборок $n_1 = n_2$, статистика рангового критерия Вилкоксона T_1 равна сумме рангов одной из выборок (любой). Чтобы применить критерий Вилкоксона, я объединил обе выборки в одной таблице, и ранжировал аккаунты по очкам (столбцы А:С).

	А	В	С	Д	Е	Ф	Г	Н
1	Очки	Выборка	Ранг	Уровень значимости	α		0,05	
2	1286	поляна	1	Объем каждой выборки			10 000	
3	1249	элита	2	Сумма рангов первой выборки	элита		80 753 951	=СУММЕСЛИ(В:В;F3;С:С)
4	1247	поляна	3	Сумма рангов второй выборки	поляна		119 256 050	=СУММЕСЛИ(В:В;F4;С:С)
5	1238	элита	4	Объем двух выборок			20 000	=G2*2
6	1237	поляна	5	Статистика	T_1		80 753 951	=G3
7	1234	поляна	6,5	Стреднее значение статистики T_1	$(T_1)_{cp.}$		100 005 000	=G2*(G5+1)/2
8	1234	поляна	6,5	Стандартное отклонение	s_{T_1}		408 258	=КОРЕНЬ(G2^2*(G5+1)/12)
9	1232	элита	8	Z-значение			-47,15	=(G6-G7)/G8
10	1228	элита	10	Уровень значимости	α		0,05	
11	1228	поляна	10		Zкр.		-1,96	=НОРМ.СТ.ОБР(G10/2)
12	1228	поляна	10					
13	1227	элита	12,5			Z < Zкр.		
14	1227	элита	12,5					
15	1226	поляна	14					
16	1224	элита	16					
17	1224	элита	16					
18	1224	поляна	16					
19	1223	элита	18,5					
20	1223	поляна	18,5					
21	1222	элита	20,5					
22	1222	элита	20,5					
23	1221	элита	22,5					
24	1221	элита	22,5					
25	1219	элита	24					
26	1218	элита	26,5					
27	1218	элита	26,5					
28	1218	элита	26,5					

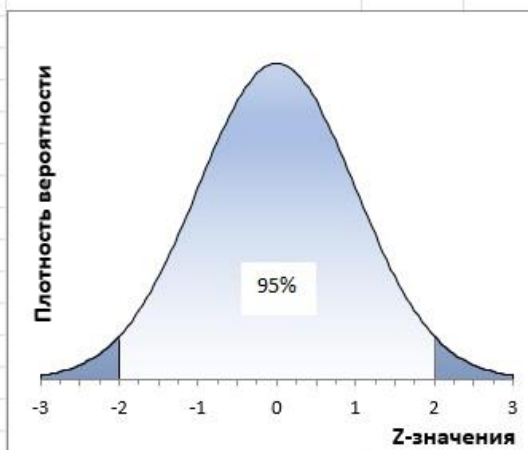


Рис. 2. Расчет критерия Вилкоксона; значения Z , попадающие в светлую область, позволяют принять нулевую гипотезу; границы равны $\pm 1,96 Z$; рассчитанное $Z = -47,15$ лежит далеко в области отклонения нулевой гипотезы

В качестве статистики T_1 я выбрал сумму рангов элиты (ячейка G6). Сумма n последовательных натуральных чисел по определению равна $n(2n+1)/2$. А выборочное среднее этой суммы задается формулой:

$$(B3) \bar{T}_1 = \frac{n_1 \cdot (n + 1)}{2}$$

где n_1 – размер одной выборки, а n – размер двух выборок; $n = 2 \cdot n_1$.

Выборочное стандартное отклонение:

$$(B4) s_{T_1} = \sqrt{\frac{n_1 \cdot n_2 \cdot (n + 1)}{12}}$$

Стандартизованная Z -статистика критерия Вилкоксона для больших выборок:

$$(B5) Z = \frac{T_1 - \bar{T}_1}{s_{T_1}}$$

Статистика Z имеет нормальное распределение.

Статистика Z для двух выборок FPL равна минус 47,15, а Z , соответствующее уровню значимости $\alpha=0,05$, минус 1,96. Z существенно меньше $Z_{кр}$. Мы отклоняем нулевую гипотезу. Медиана элиты и поляны отличаются.

Состоятельные критерии проверки однородности двух независимых выборок

Это значит, что для любых отличных друг от друга функций распределения $F(x)$ и $G(x)$ (другими словами, при справедливости альтернативной гипотезы H_1) вероятность отклонения гипотезы H_0 должна стремиться к 1 при увеличении объемов выборок m и n . Состоятельными являются критерии Смирнова и типа омега-квадрат.

Критерий Смирнова однородности двух независимых выборок

Предполагается, что функции распределения $F(x)$ и $G(x)$ непрерывны. Значение эмпирической функции распределения в точке x равно доле результатов наблюдений в выборке, меньших x . Критерий Смирнова основан на использовании эмпирических функций распределения $F_m(x)$ и $G_n(x)$, построенных по первой и второй выборкам соответственно. Значение статистики Смирнова...

$$(10) D_{m,n} = \sup_x |F_m(x) - G_n(x)|$$

...сравнивают с соответствующим критическим значением (см., например, Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.) и по результатам сравнения принимают или отклоняют гипотезу H_0 о совпадении (однородности) функций распределения. Здесь \sup – [супремум](#) – обобщение понятия максимум.

При больших объемах выборок можно воспользоваться доказанной Н. В. Смирновым в 1939 г. теоремой: в случае совпадения непрерывных функций распределения элементов двух независимых выборок

$$(11) \lim_{m \rightarrow \infty, n \rightarrow \infty} P \left\{ \sqrt{\frac{mn}{m+n}} D_{m,n} < y \right\} = K(y)$$

т. е. статистика

$$(12) S_C = \sqrt{\frac{mn}{m+n}} D_{m,n}$$

в пределе подчиняется распределению Колмогорова $K(y)$ (подробнее см. [здесь](#)).

В нашем примере функции $F(x)$ и $G(x)$ являются дискретными, но шаг составляет 1/500 диапазона, так что этим ограничением можно пренебречь. Расположим аккаунты отдельно в обеих выборках по возрастанию (столбцы А:Е рис. 3).

В столбцах G:I подсчитаем количество результатов, приходящихся на очки от 800 до 1300. В столбцах K:L подсчитаем накопленные частоты, а в столбце M – абсолютное значение разности между накопленными частотами. В ячейке P9 найдем максимум разности $|F(x) - G(x)|$. Критерий Смирнова рассчитаем по формуле (12) в ячейке P10. Критическое значение статистики Колмогорова для уровня значимости $\alpha = 0,05$ найдем из [таблицы](#) или в [работе](#).

Критерий Смирнова $S_C = 20,393$ существенно больше статистики Колмогорова $K(S) = 1,358$ при $\alpha=0,05$, гипотезу об однородности двух выборок можно уверенно отклонить. Результаты элиты и поляны отличаются.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	элита		поляна													Ячейка	Формула
2	аккаунт	очки	аккаунт	очки	x	элита	поляна	F(x)	G(x)	F(x) - G(x)						G3	=ПОСЛЕД(500;800)
3	16 409	809	4 818 840	303	800	0	7	0	0,0007	0,0007						H3	=ЧАСТОТА(B:B;G\$3#)
4	11 544	814	30 063	385	801	0	0	0	0,0007	0,0007						I3	=ЧАСТОТА(E:E;G\$3#)
5	423 487	823	1 386 823	589	802	0	0	0	0,0007	0,0007						K3	=СУММ(H\$3:H3)/10000
6	7 273	825	460 636	779	803	0	0	0	0,0007	0,0007						L3	=СУММ(I\$3:I3)/10000
7	150 155	850	4 340 012	797	804	0	0	0	0,0007	0,0007						M3	=ABS(K3-L3)
8	52 233	947	928 358	798	805	0	1	0	0,0008	0,0008							
9	4 726 349	949	3 953 627	800	806	0	0	0	0,0008	0,0008	D _{m,n}				0,2884	=МАКС(M:M)	
10	1 394	955	4 201 358	805	807	0	0	0	0,0008	0,0008	S _c				20,393	=P9*КОРЕНЬ(10000/2)	
11	11 802	956	3 982 646	808	808	0	1	0	0,0009	0,0009	уровень значимости α				0,05		
12	22 119	958	12 032	828	809	1	0	0,0001	0,0009	0,0008	K(0,05)				1,3581		
13	61 368	959	435 608	828	810	0	0	0,0001	0,0009	0,0008							
14	3 084 777	960	576 011	834	811	0	0	0,0001	0,0009	0,0008	S _c > K(0,05)					гипотеза об однородности	
15	1 622 232	963	751 357	834	812	0	0	0,0001	0,0009	0,0008						уверенно отвергается	
16	2 311 716	963	162 309	842	813	0	0	0,0001	0,0009	0,0008							

Рис. 3. Критерий Смирнова позволяет отклонить H_0

Критерий типа омега-квадрат (Лемана – Розенблатта)

Статистика критерия типа омега-квадрат для проверки однородности двух независимых выборок имеет вид:

$$(13) A = \frac{mn}{m+n} \int_{-\infty}^{+\infty} (F_m(x) - G_n(x))^2 dH_{m+n}(x)$$

где $H_{m+n}(x)$ – эмпирическая функция распределения, построенная по объединенной выборке.

$$(14) H_{m+n}(x) = \frac{m}{m+n} F_m(x) + \frac{n}{m+n} G_n(x)$$

Статистика A типа омега-квадрат зависит лишь от рангов элементов двух выборок в объединенной выборке. Данная статистика представляется в виде:

$$(15) A = \frac{1}{mn(m+n)} \left[m \sum_{i=1}^m (r_i - i)^2 + n \sum_{j=1}^n (s_j - j)^2 \right] - \frac{4mn - 1}{6(m+n)}$$

где r_i – ранг x'_i и s_j – ранг y'_j в общем вариационном ряду, построенном по объединенной выборке; здесь x'_i – элементы первой выборки x_i , переставленные в порядке возрастания.

Предел

$$(16) \lim_{m \rightarrow \infty, n \rightarrow \infty} P(A < x) = a_1(x)$$

где $a_1(x)$ – предельная функция распределения классической статистики омега-квадрат (Крамера – Мизеса – Смирнова), используемой для проверки согласия эмпирического распределения с заданным теоретическим.

Для выборок одинакового размера $n = m$ выражение (15) можно записать:

$$(ЛР1) A = \frac{1}{2n^2} \left[\sum_{i=1}^n (r_i - i)^2 + \sum_{j=1}^n (s_j - j)^2 \right] - \frac{4n^2 - 1}{12n}$$

Воспользуемся объединенным рангом двух выборок, который мы сформировали выше для расчета критерия Вилкоксона (столбцы А:С на рис. 2). Установим фильтр на столбце В – поляна, скопируем и вставим на новом месте столбцы В и С. Пронумеруем строки – см. столбцы А:С на рис. 4. Вернемся к рис. 2, установим фильтр на столбце В – элита. Повторим копипаст. Теперь в столбцах А:С на рис. 4 у нас есть объединенный ранг двух выборок (столбец В) и отдельный ранг по каждой выборке (столбец С). Осталось реализовать вычисления по формуле (ЛР1). Получим значение статистики A омега-квадрат, равное 221,9. Критическое значение a_1 для уровня значимости $\alpha = 0,05$ найдем из [таблицы](#). $A \gg a_1$, следовательно гипотезу об однородности уверенно отвергаем.

	A	B	C	D	E	F	G	H	I
1	Выборка	Ранг	Номер	(Ранг - Номер) ²					
2	поляна	1	1	0			n	10 000	
3	поляна	3	2	1			A	221,9	=(СУММЕСЛИ(A:A;"поляна";D:D)+СУММЕСЛИ(A:A;"элита";D:D))/(2*N2^2)-(4*N2^2-1)/(12*N2)
4	поляна	5	3	4	Уровень значимости		α	0,05	
5	поляна	6,5	4	6,25	Табличное значение		a_1	0,4614	
6	поляна	6,5	5	2,25					
7	поляна	10	6	16					
8	поляна	10	7	9					
10000	поляна	19999	9999	100000000					
10001	поляна	20000	10000	100000000					
10002	элита	2	1	1					
10003	элита	4	2	4					
10004	элита	8	3	25					
10005	элита	10	4	36					
10006	элита	12,5	5	56,25					
10007	элита	12,5	6	42,25					
10008	элита	16	7	81					

Рис. 4. Статистики омега-квадрат для двух выборок FPL

Когда какую-то мудреную статистику используешь впервые, нельзя быть уверенным, что всё сделал верно)) Поэтому я решил проверить результаты. Я взял две случайные выборки по 1к из 10к аккаунтов элиты, и повторил анализ омега-квадрат.

Для двух случайных выборок по 1к из элиты		
	n	1 000
	A	0,0596
Уровень значимости	α	0,05
Табличное значение	a_1	0,4614

Рис. 5. Статистики омега-квадрат для двух случайных выборок из элиты

Получился ожидаемый результат: $A \ll a_1$. Нет оснований отвергнуть нулевую гипотезу H_0 об однородности выборок.

Напомню, что элита – это отобранные 10к аккаунтов ТОП рейтинга всех менеджеров FPL по итогам пяти предыдущих сезонов. Я задался вопросом: есть ли различия на уровне значимости в результатах текущего сезона между группами ТОП-1к и ТОП-2к? Среднее значение набранных очков по ТОП-1к = 1112,3, по ТОП-2к = 1109,4. Случайно ли это различие? Визуально распределения немного отличаются:

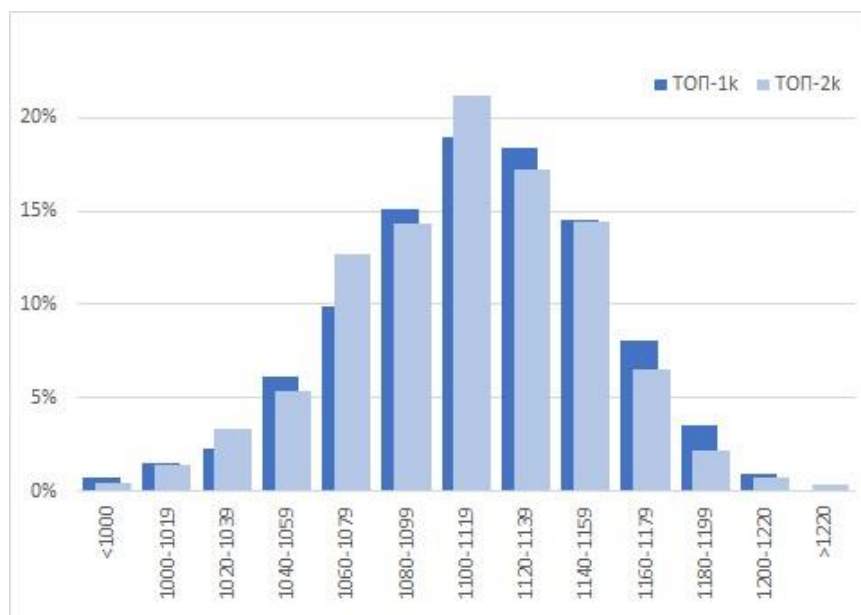


Рис. 6. Распределение текущих результатов сезона 2022/23 в группах ТОП-1к и ТОП-2к

Почти все столбцы справа от медианы более высокие у ТОП-1к. Большинство столбцов слева от медианы выше у ТОП-2к. Анализ омега-квадрат показал, что различия совсем чуть-чуть не дотянули до уровня значимости:

Для ТОП-1к и ТОП-2к элиты		
	n	1 000
	A	0,4245
Уровень значимости	α	0,05
Табличное значение	a_1	0,4614

Рис. 7. Статистики омега-квадрат первой и второй подгрупп элиты

Статистика $A = 0,4245 <$ критического $a_1 = 0,4614$. Выбери мы уровень значимости $\alpha = 0,1$, и мы смогли бы отвергнуть H_0 . Так что ситуация зыбкая...

Рекомендации по выбору критерия однородности

Мы рекомендуем для проверки однородности функций распределения (гипотеза H_0) применять статистику A типа омега-квадрат. Если методическое, табличное или программное обеспечение для статистики Лемана – Розенблатта отсутствует, рекомендуем использовать критерий Смирнова. Для проверки однородности математических ожиданий (гипотеза H'_0) целесообразно применять критерий Крамера – Уэлча. По нашему мнению, статистики Стьюдента, Вилкоксона и др. допустимо использовать лишь в отдельных частных случаях, рассмотренных выше.