ПЕРЕМЕННЫЕ ЗВЕЗДЫ

Том 12 № 3 (99) 1958

О надежности статистических работ в звездной астрономии

Н. Е. Курочкин

На основании критериев различения средних величин и дисперсий для случайных выборок устанавливается, что 1) связь, обнаруженная *О. Струве* [3] между средней абсолютной величиной лучевых скоростей V'_r и периодом P для звезд типа RR Lyr, не существует; 2) выделение подклассов по кинематике из совокупности звезд типа RR Lyr с известными V'_r, произведенное *Е. Д. Павловской* [4], оказывается ненадежным из-за большой величины дисперсии скоростей ($\sigma=100$ *км/сек*) и недостаточности материала ($n=132$). Только при значительном расхождении параметров распределения и при соответствующем расхождении других физических свойств можно с уверенностью выделять подклассы среди звезд этого типа.

On the Reliability of Statistical Investigations in Stellar Astronomy.

N. E. Kurochkin

On the basis of criterions for discerning between mean values and dispersions of random selections it has been established: 1) the relation found by *O. Struve* [3] between the mean absolute value of radial velosities V'_r and the period P for RR Lyr stars does not exist, 2) the selection of subclasses according to kinematics of RR Lyr stars with known V'_r, made by *E. D. Pavlovskaya* [4], is unreliable because of the large velocity dispersion ($\sigma=100$ km/sec) and insufficient data ($n=132$). Subclasses among stars of this type can be selected reliably only if the distribution parameters differ considerably and if there is a corresponding difference of other physical characteristics.

Многочисленность объектов, которые составляют звездные системы, позволяет судить об их свойствах на основании широкой статистической выборки. Однако к типическим характеристикам звезд по мере углубления исследований прибавляются все новые признаки. Появляется тенденция к разделению объектов на все более узкие классы. Существует нижний предел такого деления, когда ошибки наблюдений или дисперсия свойств самих объектов делают ненадежным выделение подкласса. В математической статистике на основе предельных теорем разработаны критерии, которые позволяют судить о надежности выделения подкласса или, наоборот, при заданных дисперсии и доверительной вероятности судить о минимальном объеме надежной выборки. Предельные теоремы выводятся в предположении достаточно большой выборки. В тех случаях, когда наблюдений недостаточно, применение критериев надежности дает верхние границы вероятностей или нижние границы объемов подклассов. Иными словами, при недостаточном объеме выборки становятся велики вероятности флюктуаций, и невыполнение закона больших чисел вносит свои ошибки в исследования. Все это следует учитывать при статистических работах и стремиться работать далеко за пределами случайных флюктуаций.

Некоторая величина, подверженная случайным колебаниям, характеризуется своим средним значением и дисперсией в данной выборке.

Назовем генеральной такую выборку, свойства которой мы можем считать идентичными действительности. Тогда расхождение между средней нашей выборки и средней генеральной совокупности характеризуется вероятностью [1]:

$$P(t) = P\left\{|a - \bar{x}| \leqslant \frac{t\sigma}{\sqrt{n}}\right\} = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{z^2}{2}} dz. \qquad (1)$$

Здесь a и $\bar{x}$ — средние генеральной совокупности и выборки из нее, t — параметр, характеризующий доверительную вероятность $P(t)$, σ — дисперсия x_i, n — объем выборки. Из формулы (1) устанавливается, что при заданной доверительной вероятности и параметре t ($P(t)$ находится по таблицам, как функция t) отклонение выборочной средней от средней генеральной совокупности (или приближенно — двух выборочных средних) с вероятностью $P(t)$ оказывается не больше величины

$$\Delta = |a - \bar{x}| = \frac{t\sigma}{\sqrt{n}}. \qquad (2)$$

При выделении подкласса в качестве генеральной совокупности мы можем взять объем всего класса.

Для статистических работ можно считать приемлемой величину, соответствующую $t = 1.00$, или $P(t) = 0.68$. В более строгих исследованиях следует принимать $t = 2.0$ ($P(t) = 0.95$) и $t = 3.0$ ($P(t) = 0.997$) или в каждом случае оценивать вероятность неслучайности расхождения средних. В табл. 1 даются те минимальные объемы выборок, которые необходимы, чтобы с вероятностью 0.68 фиксировать расхождение средних Δ при данной дисперсии σ.

Таблица 1

Наименьший надежный объем выборки при заданных дисперсии σ и расхождении средних Δ (при $P(t) = 0.68$)

Δ	σ							
	5	10	20	50	75	100	150	200
5	1	4	16	100	225	400	900	1600
10		1	4	25	57	100	225	400
15			2	12	25	45	100	178
20			1	7	15	25	57	100
25				4	9	16	36	64
30				3	7	12	25	45
40				2	4	7	14	25
50				1	3	4	9	16
75					1	2	4	8
100						1	3	4

Если мы имеем дисперсии s_1 и σ случайной выборки и генеральной совокупности, то вероятность ω, что σ заключено между

$$\frac{\sqrt{n}\, s_1}{t_1} \leqslant \sigma < \frac{\sqrt{n}\, s_1}{t_2} \tag{3}$$

равна

$$\omega = \frac{1}{2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right)} \int_{t_1}^{t_2} t^{(n-3)} e^{-\frac{t^2}{2}} dt. \tag{4}$$

Таблицы функций $\omega(t,k)$ даются в [1]. Оценка расхождения двух выборочных дисперсий может основываться на отношении их квадратов $T_k = \frac{\sigma^2_1}{\sigma^2_2}$ $(\sigma_1 > \sigma_2)$. В практически очень удобных таблицах в книге *Романовского* [2] даются верхние пределы $T_k = \frac{\sigma^2_1}{\sigma^2_2}$ $(\sigma_1 > \sigma_2)$, при которых расхождение дисперсий двух выборок еще случайно (при доверительных вероятностях 0.05 и 0.01). По табл. 2 можно определить наименьший объем выборки при заданном расхождении средних, в качестве σ_1 можно взять максимальное значение дисперсии из многих выборок.

Таблица 2

Наименьший надежный объем выборки при заданном расхождении дисперсий

$\frac{\sigma^2_1}{\sigma^2_2}$	$1-P(t)$ 0.05	0.01	$\frac{\sigma^2_1}{\sigma^2_2}$	$1-P(t)$ 0.05	0.01
1.00	∞	∞	1.75	23	45
1.10	300	430	2.00	16	30
1.15	188	275	2.50	10	19
1.25	106	166	3.00	8	14
1.50	42	75			

Мы применим изложенные сведения к некоторым закономерностям, обнаруженным в движениях короткопериодических цефеид.

В 1950 г. *Струве* [3] нашел зависимость между средней величиной лучевых скоростей $|\overline{V_r}|$ и периодом у звезд типа RR Lyr. В табл. 3 приведена эта зависимость по работе *Е. Д. Павловской* [4], которая ревизовала данные *Струве*, в частности исключила из лучевых скоростей цефеид движение Солнца. В табл. 3 указаны: пределы периодов, $\overline{P}$—средний период для группы, $|\overline{V'_r}|$ — средняя лучевая скорость, n— число звезд, отнесенных к группе с данными пределами периодов, $n_{\min}$—минимальный объем выборки, который необходим, чтобы при дисперсии $\sigma = 100$ *км/сек* отклонение от средней скорости для всех этих звезд $|\overline{V'_r}| = 76$ *км/сек* фиксировалось как неслучайное с вероятностью 0.68. Мы видим, что n меньше $n_{\min}$ почти во всех случаях, что показывает ненадежность выведенной зависимости между $|\overline{V_r}|$ и P.

Наблюдающиеся уклонения средних для выделенных подклассов с большой долей вероятности могут быть случайными.

Таблица 3

Пределы P	$\overline{P}$	$\overline{\lvert V'_r\rvert}$, *км/сек*	n	$n_{\min}$
$P<{}^{d}.30$	0.196	46.4±10.4	17	12
0.30—0.40	0.352	60.4 9.1	20	41
0.40—0.50	0.458	75.2 13.6	30	>400
0.50-0.65	0.559	89.5 12.0	49	71
$0.65<P$	0.772	56.8 10.9	16	29

Оправдывание закономерности, установленной по трем точкам, для четвертой точки, повышает вероятность зависимости. Если по трем точкам мы считаем вероятность связи равной 0.68, то для четырех точек вероятность связи повышается до $[1-(1-0.68)(1-0.68)]=0.90$. Наоборот, выпадение одной точки из закономерности уменьшает надежность. Оба эти случая имеются в табл. 3.

Таблица 4

Пределы P	n	$\overline{V'_r}$	$n_{\min}$	σ, *км/сек*	$T_{\text{набл}}$		T_k при $[1-P(t)]$ 0.05	0.01
$P<0^{d}.25$	9	−1	63	86.6	2.25	1.40	2.93	4.86
0.25—0.35	19	+7	45	67.2	3.73	2.34	1.92	2.57
0.35—0.45	19	+10	45	65.0	3.96	2.50	1.92	2.57
0.45—0.55	41	+41	30	129.3	1.00	—	1.52	1.81
0.55—0.65	28	−13	36	102.8	1.59	1.00	1.67	2.10
$0.65<P$	14	+1	51	73.7	3.10	1.95	2.21	3.16
Все звезды	130	+12		96.2	1.81	1.44	1.60	1.90

При нормальном распределении V_r' изменение $\lvert\overline{V_r'}\rvert$ в выделенных подклассах может быть связано или с изменением нульпункта или с изменением дисперсии σ. В табл. 4 приведены средние $\overline{V_r'}$ и дисперсии σ для групп звезд с различными периодами. Ясно видно, что систематический ход $\overline{V_r'}$ и σ в зависимости от периодов отсутствует. Отношение квадратов дисперсий оказывается вблизи границы случайности. Дисперсия при малых объемах выборки весьма чувствительна к флюктуациям, как показывает табл. 5, где сравниваются случайные выборки. В ряде случаев для случайных выборок мы получаем существенные расхождения дисперсий, что, несомненно, связано только с малым объемом выборки.

Таблица 5

n	σ, *км/сек*	$T_{\text{набл}}$		T_k при $[1-P(t)]$ 0.05	0.01
5	101	1.15	—	5.63	13.46
10	59	3.37	2.88	2.71	4.31
20	95	1.27	1.11	1.88	2.49
30	78	1.90	1.64	1.64	2.03
34*	73	2.19	1.88	1.59	1.95
98*	108	1.00	—	1.26	1.40
132**	100	1.17	1.00	1.26	1.39

* Группа, выделенная *Е. Д. Павловской*.
** Все звезды.

Критерий для сравнения средних несколько выделяет группу звезд с периодами $0.^d45$—$0.^d55$. Здесь наблюдаются, как показывает рис. 1, несколько звезд с аномально большими скоростями V_r', причем 11 звезд с $V_r' > +100$ и 6 звезд с $V_r' < -100$; 28 звезд с $V_r' > 0$ и 13 звезд $V_r' < 0$. По дисперсии, как показывает критерий T_k (ср. табл. 4, 7-й столбец) эта группа, по-видимому, также отличается от других групп (хотя различие от дисперсии всей совокупности находится на границе случайности). Можно думать, что расхождение средних и дисперсий

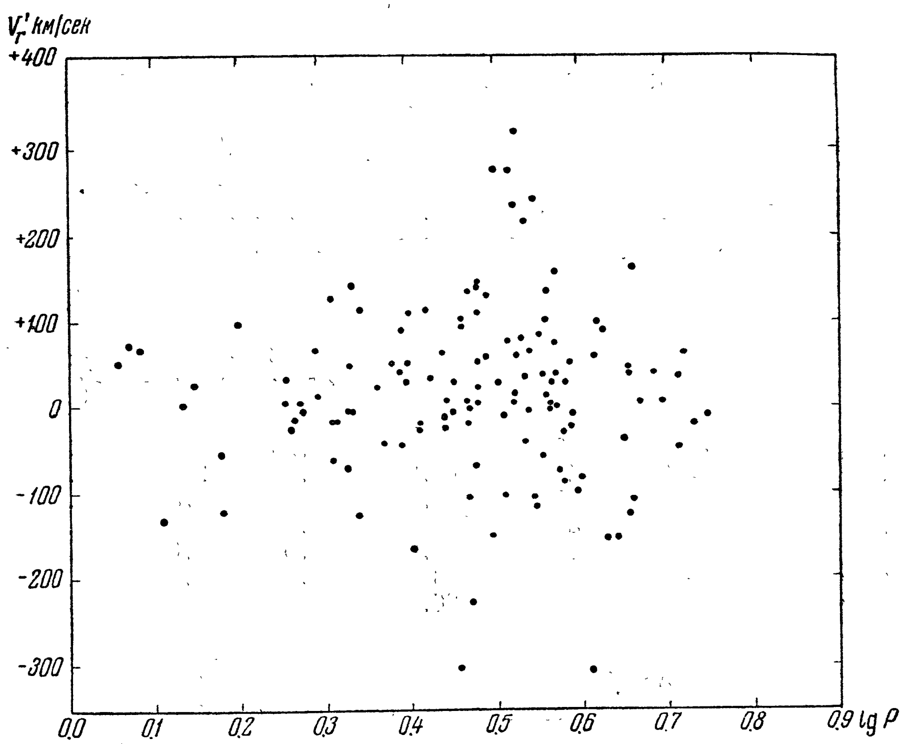

Рис. 1

здесь также случайно. Это подтверждает вычисление расхождения Δ по формуле (2) при более строгих критериях $t=2.0$ и $P(t)=0.95$. При этих условиях $\Delta=30$ *км/сек*, тогда как расхождение в нашем случае составляет 29 *км/сек* (от генеральной средней при $n=130$). Исследование пространственных скоростей короткопериодических цефеид обнаруживает совсем иные закономерности, причем по большим величинам компонент скоростей выделяются звезды с обычными V_r'. Отсутствие систематического хода V_r' и σ в зависимости от периода приводит в выводу, что зависимость, обнаруженная *Струве*, в действительности не имеет места. Это непосредственно видно из рис. 1. Возможно, имеется реальный избыток звезд с положительными лучевыми скоростями.

Из изложенного видно, с какой осторожностью следует подходить к изучению статистических данных при недостаточном материале. Ограниченность всякой статистической выборки всегда приводит к занижению дисперсии, так как „хвосты“ нормального распределения отсекаются. В нашем случае это неизбежно вызовет зависимость между средней абсолютной величиной $|\overline{V_r'}|$ и объемом выборки. На рис. 2 показана такая статистическая связь между $|\overline{V_r'}|$ и n. Очевидно, что зависимость $|\overline{V_r'}|$ и n не менее определенна, чем зависимость между $|\overline{V_r'}|$ и $\overline{P}$. Это следствие невыполнения основного условия статистических работ—закона больших чисел.

В заключение остановимся на другом выводе работы Е. Д. *Павловской*: о разделении короткопериодических цефеид на две группы по физическим характеристикам. Малое количество звезд ($n=132$) требует здесь, как мы установили, весьма строгого анализа.

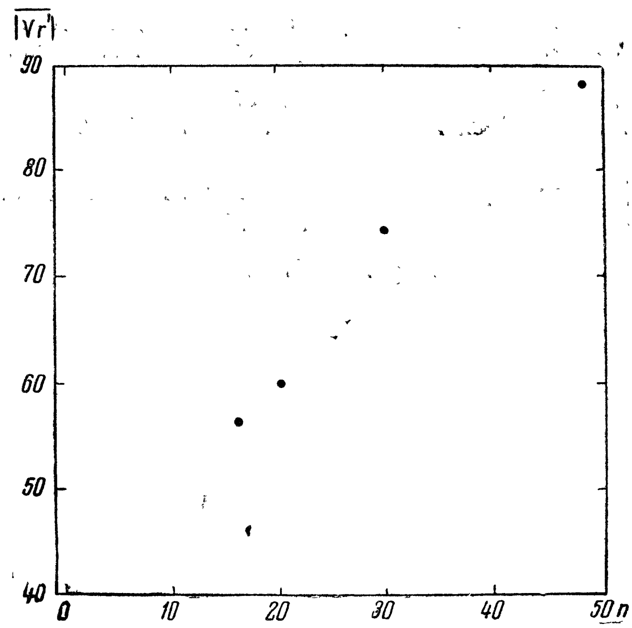

Рис. 2

Разделение на неравные группы создает указанную выше опасность занижения дисперсии для малочисленных групп.

В табл. 5 проводится сравнение групп, выделенных *Е. Д. Павловской*, с выборками, произведенными по таблицам случайных чисел [2] (из пространственных скоростей).

Применен описанный в [2] критерий разделения дисперсий по их отношению T_k.

Во всех случаях различие дисперсий оказывается вблизи и ниже границы случайных расхождений. Это означает, что свойства произведенных выборок из нашей совокупности звезд типа RR Lyr различаются несущественно. Всякая попытка разделить звезды на еще более мелкие группы по более тонким физическим признакам приведет к бо̀льшим трудностям.

Какие-либо надежные выводы можно получить только при сопоставлении физических свойств с разными независимыми кинематическими характеристиками.

Укажем в заключение на некоторое общее значение рассмотренного здесь вопроса. До тех пор, пока совокупность признаков не имеет объяснения, т. е. пока свойства этой совокупности не поставлены в связь с известными общими физическими законами, наблюдаемые факты имеют чисто феноменологическое значение. Они рассматриваются статистически. Учет надежности в этих случаях оказывается совершенно необходимым. При изучении признаков, характеризуемых качественно, вывод также может быть недостаточно надежен, если „дисперсия" характеристик значительна.

Не во всех исследованиях можно количественно оценить расхождение исходных данных. Это следует иметь в виду при оценке качественных суждений.

Литература

1. *Б. В. Гнеденко,* Курс теории вероятностей. М., 1954.
2. *В. И. Романовский,* Применения математической статистики в опытном деле. М.—Л., 1947.
3. *O. Struve,* PASP **62**, 217, 1950.
4. *Е. Д. Павловская,* ПЗ **9**, № 6, 349, 1953.

Государственный астрономический институт
им. Штернберга

Апрель 1957 г.