

УДК 517

СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ПО
КОЛИЧЕСТВЕННОМУ ПРИЗНАКУ

Артемьев Н.А., студент гр. СРБ-16, 1 курс

Научный руководитель: Романовский С.А., преподаватель

Федеральное казенное профессиональное образовательное учреждение
«Новокузнецкий государственный гуманитарно-технический колледж-
интернат» Министерства труда и социальной защиты Российской Федерации,
г. Новокузнецк

Актуальностью исследования является то, что статистические методы, являются эффективным инструментом сбора и анализа информации по количественному признаку, например, о качестве. Применение этих методов, не требует больших затрат и позволяет с заданной степенью точности и достоверностью судить о состоянии исследуемых явлений (объектов, процессов).

Цель исследования состоит в статистической обработке результатов выборочного контроля партии M&M's с арахисом. Объем выборки составил 150 штук. В качестве параметра качества выбрана длина драже. Измерения производились штангенциркулем с ценой деления 0,05 мм.

Задачи исследования:

- измерить драже;
- произвести статистическую обработку результатов;
- проверить согласия опытного распределения с теоретическим нормальным.

Делаем выборку $n=150$ штук из генеральной совокупности и измеряем выбранный параметр каждого объекта.

Подготовим полученные данные для дальнейшей статистической обработки. Для этого представим их в виде упорядоченного ряда путём составления контрольного листа.

Так как много данных, то это затрудняет их дальнейший анализ и математическую обработку для определения числовых характеристик ряда. Поэтому целесообразно представить данные в виде интервального ряда.

Для этого весь диапазон измеренных значений разобьём на ряд интервалов, подсчитывая при этом количество измерений, попавших в каждый интервал. Определяем количество интервалов k по формуле

$$k = \sqrt{n} = \sqrt{150} = 12,25 \approx 12, \quad (1)$$

где n – объём выборки [1, с. 9].

При таком выборе числа интервалов среднюю арифметическую и дисперсию можно вычислить достаточно точно, объём вычислений при этом будет не слишком велик.

Определим ширину каждого интервала

$$\frac{(x_{max}-x_{min})}{k} = \frac{23,05-13,35}{12} = 0,808 \approx 0,81, \quad (2)$$

где x_{max} – наибольшее значение среди измеренных в выборке,

x_{min} – наименьшее значение среди измеренных в выборке,

k – количество интервалов.

Далее составим таблицу интервального ряда. В каждый интервал будем включать те наблюдения, числовые значения которых больше или равны нижней границе интервала и меньше верхней.

При статистическом анализе закономерностей распределения экспериментальных данных используют обычно две группы числовых характеристик, которые в сжатой форме характеризуют результат исследований.

Одна из них описывает среднее положение наблюдаемых значений. Из этой группы характеристик чаще всего используют среднее арифметическое случайной величины и медиану.

Вторая группа числовых характеристик описывает рассеяние единичных значений случайной величины от её среднего значения. Сюда относят дисперсию, стандартное отклонение и размах.

Рассчитаем данные характеристики.

Характеристики положения:

- Средняя арифметическая

$$\bar{x} = \sum_{i=1}^k x_i^{cp} \cdot h_i = \frac{2663,46}{150} = 17,76 \quad (3)$$

где $\bar{x}$ – средняя арифметическая,

x_i^{cp} – середина i -го интервала,

h_i – число попавших в i -ый интервал значений [3].

- Медиана

Медиана равна при чётном числе измеренных значений полусумме двух измеренных значений, расположенных в середине упорядоченного ряда [3]:

$$M_e = \frac{x_i + x_{i+1}}{2} = \frac{x_{75} + x_{76}}{2} = \frac{17,75 + 17,85}{2} = 17,8 \quad (4)$$

Характеристики рассеивания:

- Размах [3]

$$R = x_{max} - x_{min} = 23,05 - 13,35 = 9,7. \quad (5)$$

- Дисперсия [3]

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^k (x_i^{cp} - \bar{x})^2 \cdot h_i = \frac{485,815806}{150} = 3,24 \quad (6)$$

где σ^2 – дисперсия,

n – объём выборки,

x_i^{cp} – среднее значение интервала,

$\bar{x}$ – средняя арифметическая,

h_i – число попавших в i -ый интервал значений [3].

- Стандартное отклонение [3]

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n} \sum_{i=1}^k (x_i^{cp} - \bar{x})^2 \cdot h_i} = \sqrt{3,24} = 1,8 \quad (7)$$

Рис 1. Графическое представление интервального ряда



Гистограмма не имеет ярко выраженного пика, на ней можно увидеть плато. Это может являться следствием того, что в совокупности исследуемых данных могут быть смешаны результаты нескольких процессов. В целом можно предположить, что распределение подчиняется нормальному закону.

Очень часто измеряемые численные параметры, являющиеся случайными величинами, подчиняются нормальному закону распределения. Поэтому при планировании и проведении измерений по количественному признаку полагают, что контролируемый параметр имеет точно или приближенно нормальное распределение.

В случае отсутствия такой уверенности целесообразно осуществить проверку согласия опытного распределения контролируемого параметра с нормальным законом.

Такую проверку обычно проводят с использованием критерия согласия Пирсона или критерия χ^2 («хи-квадрат»).

Этот критерий согласия позволяет провести проверку гипотезы о соответствии опытного распределения случайной величины теоретическому.

Для проверки с помощью критерия χ^2 нормальности опытного распределения рассчитаем меру расхождения опытного и нормального теоретического распределений

$$Y = \sum_{i=1}^k \frac{(h_i - np_i)^2}{np_i}, \quad (8)$$

Где np_i – теоретическая частота i -го интервала значений x ,

p_i – вероятность того, что наблюдаемая случайная величина x_i генеральной совокупности попадает в i -ый интервал [4, с. 5].

Для расчёта значений p_i необходимо перейти к нормированной случайной величине U [4, с. 5]:

$$U_i = \frac{(x_i - \frac{x_i^{cp} \cdot h_i}{n})}{\sigma} \quad (9)$$

$$F(U_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} - \text{плотность стандартного нормального распределения} \quad (10)$$

Таблица 5 – Расчёт критерия согласия Пирсона

Интервал $x_{i+1} - x_i$	h_i	U_i	$F(U_i)$	$p_i = F(U_i) - F(U_{i-1})$	np_i	$(h_i - np_i)$	$(h_i - np_i)^2$	$\frac{(h_i - np_i)^2}{np_i}$
13,35-14,16	1	-						
		1,998379	0,022838	0,022838	3,425671			
14,16-14,97	9	-				0,883625	0,780793	0,085647
		1,548293	0,060776	0,037938	5,690704			
14,97-15,78	11	-			11,29214	-		
		1,098208	0,136057	0,075281	6	0,292146	0,085349	0,007558
15,78-16,59	22	-			18,35939		13,25402	
		0,648123	0,258453	0,122396	2	3,640608	8	0,721921
16,59-17,40	21	-			24,45826	-	11,95959	
		0,198038	0,421508	0,163055	4	3,458264	2	0,488980
17,40-18,21	22	-			26,69851	-	22,07601	
		0,252048	0,599498	0,177990	2	4,698512	4	0,826863
18,21-19,02	27	-			23,88059			
		0,702133	0,758702	0,159204	7	3,119403	9,730675	0,407472
19,02-19,83	21	-			17,50236		12,23347	
		1,152218	0,875384	0,116682	2	3,497638	0	0,698961
19,83-20,64	9	-			10,51071	-		
		1,602304	0,945456	0,070071	6	1,510716	2,282262	0,217137
20,64-21,45	4	-			5,171745			
		2,052389	0,979934	0,034478	5,171745			
21,45-22,26	0	-			2,084926	-		
		2,502474	0,993834	0,013900	2,084926	0,945272	0,893540	0,112462
22,26-23,07	3	-			0,688601			
		2,952560	0,998424	0,004591	0,688601			
							Всего	3,567001

Таким образом, критерий Пирсона равен $Y=3,567001$.

Рассчитаем критическое значение $\chi^2_{\alpha, m}$.

Число степеней свободы m равно:

$$m = k - 3. \quad (11)$$

Число интервалов $k = 9$. Значит $m = 9 - 3 = 6$.

Уровень значимости $\alpha = 0,05$.

Используя функцию в электронных таблицах «Excel» ХИ2ОБР находим критическое значение:

$$\chi^2_{\alpha, m} = \chi^2_{0,05,6} = 12,591587.$$

Сравним критическое значение и критерий Пирсона [4, с. 6]:

$$12,591587 > 3,567001$$

$$\chi^2_{0,05,6} > Y.$$

Исходя из полученного неравенства можно сделать вывод, что генеральная совокупность подчиняется нормальному закону распределения при выбранной вероятности ошибки $\alpha=0,05$.

Вывод

Результаты измерений представлены в виде упорядоченного и интервального рядов. Рассчитаны характеристики положения (средняя арифметическая, медиана) и рассеяния (дисперсия, стандартное отклонение и размах). Построенная по результатам измерения гистограмма не имеет ярко выраженных пиков и имеет плато, что говорит о существовании нескольких процессов, формирующих рассматриваемую партию драже (генеральную совокупность).

Выдвинутая гипотеза о нормальном распределении в ходе ее проверки при помощи критерия Пирсона была принята, т.е. распределение случайных величин в генеральной совокупности подчиняется нормальному закону распределения.

Результаты исследования показали, что генеральная совокупность подчиняется нормальному закону распределения, процесс изготовления драже стабилен.

Результаты подтверждают утверждение о том, что измеряемые численные параметры продукции, являющиеся случайными величинами, чаще всего на практике подчиняются нормальному закону распределения.

Список литературы:

1. Сильвестров Ю.Г. Статистическая обработка результатов выборочного контроля по количественному признаку: метод. указ. / Сиб. гос. индустр. ин-т.; сост. Ю.Г. Сильвестров. – Новокузнецк: Изд. центр СибГИУ, 2014. – 36 с.
2. ДП СМК 4.2.3-3.0-2009. Управление документацией. Оформление выпускных квалификационных работ, отчетов по практике, курсовых проектов и работ / СибГИУ. – Введ. 26.10.2009. – Новокузнецк: Изд. центр СибГИУ, 2009. – 33 с.
3. Строителев В.Н. Статистические методы в управлении качеством: учебник для вузов / В.Н. Строителев, В.Е. Яницкий. - М.: Европейский центр по качеству, 2002. - 164 с.
4. Сильвестров Ю.Г. Проверка гипотезы о законе распределения генеральной совокупности: метод. указ. / Сиб. гос. индустр. ин-т.; сост. Ю.Г. Сильвестров. – Новокузнецк: Изд. центр СибГИУ, 2014. – 20 с.