«Статистическая проверка гипотез»

Шарифуллин Ринас Рамилевич гр. 09-131

Вариант – 13

Задание 3_1

1. В соответствии с молекулярной теорией при термической обработке металла должно происходить изменение его прочности (увеличение или уменьшение).
2. Для проверки предположения были произведены измерения прочности до и после обработки n = 78 металлических прутков, изготовленных из одной плавки металла.
3. По соображениям физики процесса можно предположить, что измерения суть реализации нормальной случайной величины N (μ, σ^2), где μ – математическое ожидание показателя прочности в образцах, σ^2 – дисперсия, характеризующая степень изменчивости этого показателя от образца к образцу, а также точность метода измерения. До проведения обработки $\mu = \mu_1$, $\sigma^2 = \sigma_1{}^2$; после проведения обработки $\mu = \mu_2$, $\sigma^2 = \sigma_2{}^2$.
4. Ожидается, что $\theta = \mu_1 - \mu_2 \neq 0$, т.е. в среднем прочность металла изменяется. Нулевая гипотеза H_0: $\theta = 0$ при альтернативе H_1: $\theta \neq 0$.
5. Уровень значимости $\alpha = 0.01$.
6. Ввиду предположения нормальности наблюдений следует применить одновыборочный (разностный) критерий Стьюдента, основанный на разности $\bar{x}_1 - \bar{x}_2$ выборочных средних значений до и после обработки. Для вычисления статистики Стьюдента необходимо найти среднее арифметическое $\bar{u}$ и дисперсию (смещённую) $S_u{}^2$ разностей $u_i = x_i - y_i$, $i = 1, \ldots, n$. Статистика Стьюдента равна
$$T = \frac{\bar{u}}{\sqrt{S_u^2}}\sqrt{n-1}$$
В соответствии с теоретическими предпосылками ожидается, что абсолютная величина $|\bar{x}_1 - \bar{x}_2|$ будет принимать большие положительные значения. Нулевая гипотеза должна отвергаться, когда абсолютное значение статистики Стьюдента $|T| \geq C$.
7. Функция распределение тестовой статистики в граничной точке $\theta_0 = 0$ совпадает с функцией распределения Стьюдента $\mathbb{St}_{(n-1)}$ с $n - 1 = 77$ степенями свободы.
8. Критическая константа C_α находится из уравнения

$$P\{|T| > C\} = 2(1 - \mathbb{St}_{77}(C)) = 0.01,$$

т.е. – равна верхней 0.005-квантили распределения Стьюдента с 77 степенями свободы.

a) Воспользовавшись таблицей (пакетом Excel), нашли, что $C_\alpha = 2{,}641$.
b) Окончательный вид критической области $|T| \geq 2{,}641$

9.

a) По представленным данным найдено

		До	После	По разностям
Объём выборки	n	78	78	78
Среднее	$\bar{x}$	108,07	105,11	2,96
Станд. откл.	s	8,133	9,21	12,654
Станд. ошибка среднего	m	0,921	1,04	1,43
Статистика Стьюдента			$T = 2{,}05 = t_{эксп}$	
1%-ая критическая область			$\|T\| \geq 2{,}641$	
Гипотеза отсутствия эффекта			принимается	
с критическим уровнем значимости			p-val = 0,043	
Вывод. Отклонение от нулевой гипотезы статически не значимо.				

b) Критический уровень значимости p-value вычислялся по формуле

$$\text{p-val} = 2(1 - \mathbb{St}_{77}(|T|)) = 2(1 - \mathbb{St}_{77}(2{,}05)) = 0{,}043.$$

Наблюдения не противоречат гипотезе отсутствия эффекта обработки (П-значение p-val = 0,043).

Задание 3_2

1. Требуется сравнить точность измерений, производимых двумя хорошо откалиброванными (т.е. без систематической ошибки) приборами. Первый прибор был изготовлен по новой технологии, которая, как заявляют изобретатели, повышает точность.
2. Первым прибором было произведено 64 измерений эталонных образцов, вторым прибором – 45 измерений.
3. Можно предположить, что ошибка измерения каждым из приборов носит случайный характер и имеет нормальное распределение со средним нуль и дисперсиями $\sigma_1{}^2$ и $\sigma_2{}^2$.
4. Ожидается, что $\sigma_1{}^2 < \sigma_2{}^2$. Т.е. в терминах параметра $\theta = \frac{\sigma_1^2}{\sigma_2^2}$ нулевая гипотеза $\mathbf{H}_0$: $\theta \geq 1$ при альтернативе $\mathbf{H}1$: $\theta < 1$.
5. Уровень значимости α = 0.05.
6. В силу нормальности распределения наблюдений, можно применить критерий Фишера. Тестовая статистика Фишера
$$F = \frac{\tilde{s}_1^2}{\tilde{s}_2^2},$$
где $\tilde{S}_j{}^2$ – несмещённая оценка дисперсии в j-й группе (j-го прибора), j = 1,2. Ожидания разработчиков нового прибора будут подтверждены, если F примет достаточно большие значения, т.е. критическая область имеет вид $\{F > C\}$.
7. В граничной точке θ_0=1 распределение статистики Фишера совпадает с распределением Фишера $\mathbb{F}_{(n1-1,n2-1)}$ с n_1 = 64 и n_2 = 45 степенями свободы.
8. Критическая константа C_α находится как решение уравнения
$$P\{F > C_\alpha\} = 1 - \mathbb{F}_{(64,45)}(C) = 0.05,$$
т.е. равна верхней 0.05-квантили распределения Фишера.
   a) По таблице распределения Фишера (процедуре Excel), находим C_α = 1,596.
   b) Окончательный вид критической области $\{F > 1{,}596\}$.
9. По представленным данным:

	1-ый прибор	2-ый прибор
n	64	45
$\bar{x}$	128,69	149,149
$\tilde{s}^2$	59,898	182,537
Статистика Фишера $F = s_1^2 \,/\, s_2^2$		3,047
5%-ая критическая область		$F \geq 1{,}596$
Гипотеза H_0: $\sigma_1^2 \geq \sigma_2^2$		отвергается

Вывод: предположение о повышенной точности 1-го прибора статистически подтверждается	
p-значение	0,00007

a. p-value вычисляется по формуле

$$\text{p-val} = 1 - \mathbb{F}_{(64,45)}(3{,}047) = 0{,}00007$$

Можно сделать вывод о высокой значимости согласия данных (П-значение p = 0.00007) с претензиями разработчиков нового прибора.

Задание 3_3

1. В регионе 0.4 населения ежегодно страдало от ОРЗ. Фармацевтическая компания разработала средство профилактики ОРЗ и обещала, что это средство приведёт к понижению доли заболевших.
2. Для проверки этого заявления предполагается разработанное средство применить к группе n = 68 случайно отобранных пациентов.
3. Таким образом, в эксперименте наблюдаются бернуллиевские случайные величины с вероятностью успеха (не подхватить ОРЗ) θ.
4. Ожидается, что θ > 0,6. Нулевая гипотеза $\mathbf{H}_0$: $\theta \le 0{,}6$ при альтернативе $\mathbf{H}_1$: $\theta > 0{,}6$.
5. Уровень значимости α = 0.05.
6. Применим критерий знаков, основанный на числе T, избежавших заболевания после вакцинации. Ясно, что ожидания фармкомпании будут подтверждены, если T примет достаточно большое значение, т.е. критическая область имеет вид $\{T \ge C\}$.

В граничной точке $\theta_0 = 0{,}6$ функция распределения статистики T есть функция биномиального распределения $\mathbb{Bim}(k \mid n, \theta_0) = \mathbf{P}_{\theta 0}\{T \le k\}$ с n = 68, $k = 0{,}1, \dots, n$.

7. Критическая константа C_α находится как решение неравенства
$$\mathbf{P}\{T \ge C_\alpha\} = 1 - \mathbb{Bim}(C_\alpha - 1 \mid 68, 0.6) \le 0.05,$$
причём, из всех таких констант нужно выбрать наименьшую, т.е. число $C_\alpha - 1$ равно квантили порядка 0.95 биномиального распределения.
    a. По таблице биномиального распределения (с помощью процедуры Excel), находим C_α = 48.
    b. Вид критической области: нулевая гипотеза отвергается, если $\{T \ge 48\}$.
8. По представленным данным:

Частота появления А (не заболевших)	0,603 41 из 68
5%-ая критическая область	≥ 48
Гипотеза **H₀**: $p \le 0{,}6$	принимается
<u>Вывод.</u> Отклонение от нулевой гипотезы статистически не значимо. Имеются все основания не одобрять применение профилактического средства.	
Критический уровень значимости	$\alpha_{crit} = 0{,}53$

   a. p-value вычисляется по формуле

$$\text{p-val} = 1 - \mathbb{Bin}(41 - 1 |\ 68; 0{,}6) = 0{,}53$$

Можно сделать вывод о не согласии данных (p > 0,05) с ожиданиями фармкомпании.

Задание 3_4

1. Для увеличения срока службы электроламп был разработан новый дизайн цоколя.
2. Чтобы проверить действенность этой модификации предлагается провести испытания на долговечность в одинаковых условиях партии m = 64 ламп, изготовленных со старым цоколем, и, независимо, партии n = 48 ламп с новым цоколем.
3. Время службы каждой лампы есть случайная величина с функцией распределения F_1 (для старых образцов – 1-я выборка) или F_2 (для новых образцов – 2-я выборка).
4. Ожидается, что $F_1(x)$ > $F_2(x)$ для $\forall x$ > 0 (т.е. для ламп старого образца более вероятен выход из строя до момента x или, другими словами, $\xi_2 \succ (d)\ \xi_1$ – время службы новых ламп стохастически больше). Гипотеза $\mathbf{H}_0$: $F_1(x) \equiv F_2(x)$ при альтернативе $\mathbf{H}_1$: $F_1(x) > F_2(x), \forall x$.
5. Уровень значимости α = 0.025.
6. Применим критерий Вилкоксона, основанный на сумме W рангов 1-й выборки в общем ряду данных. Если справедлива альтернативная гипотеза (наблюдения в 1-й выборке стохастически меньше наблюдений во 2-й), то ожидаются небольшие значения W. Другими словами, критическая область имеет вид $\{W \le C\}$.
7. При справедливости нулевой гипотезы распределение статистики W есть распределение Уилкоксона с параметрами (64, 48). Можно применить нормальную аппроксимацию с матем.ожиданием μ_W = 3616 и стандартным отклонением σ_W = 170,08.
8. Т.о., критическая константа C_α находится как целая часть решения уравнения

$$P\{W \le C\} \approx \Phi\left(\frac{C-3616}{170{,}08}\right) = 0{,}025,$$

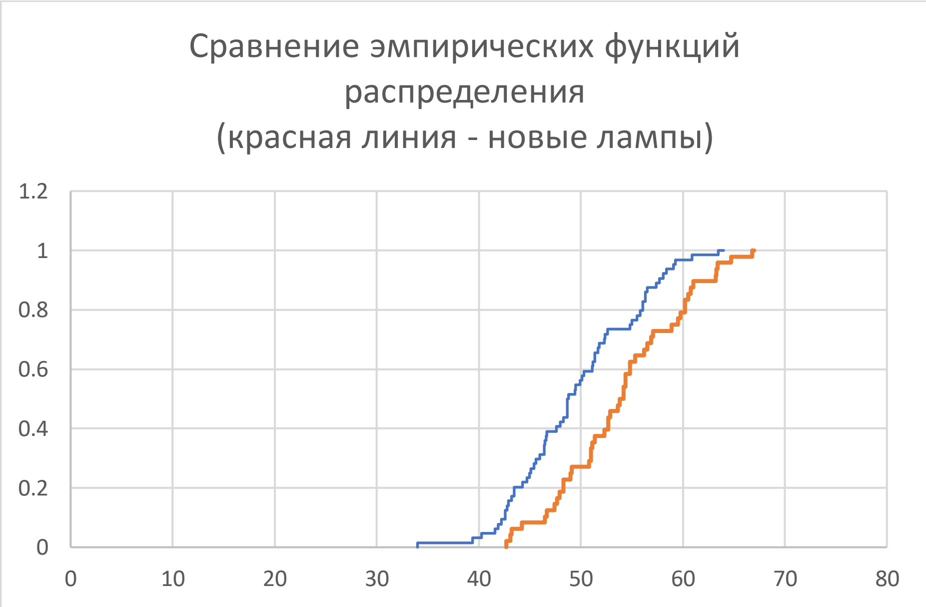
т.е. C равна квантили порядка 0.05 нормального закона.

   a. По таблице (с помощью процедуры Excel), находим C_α = 3282,649.
   b. Вид критической области: нулевая гипотеза отвергается, если $\{W \le 3282{,}649\}$.

9. По представленным данным:

Объемы выборок	m = 64	N = 48
Сумма рангов 1-ый выборки W	3028	

Математическое ожидание μ_w		3616	
Стандартное отклонение σ_w		170,08	
2,5%-я критическая область		$W \leq 3282{,}649$	
Вывод	Нулевая гипотеза о совпадении распределений		отвергается
	с критическим уровнем значимости	$\alpha_{crit} = 0{,}00027$	
<u>Заключение.</u> Новый дизайн цоколя лампы приводит к большому увеличению срока службы.			



a. p-value вычисляется по формуле

$$\text{p-val} \approx \Phi\left(\frac{3028-3616}{170{,}08}\right) = 0{,}00027$$

Поскольку p-val значительно меньше 2,5%-го уровня значимости, то есть все основания считать новый цоколь более надёжным.

Задание 3_5

1. Производители кока-колы уверяют, что содержание витамина С в банках идентично содержанию этого витамина в стеклянных бутылках.
2. Измерено содержание витаминов в n_1 = 75 банках (группа A) и в n_2 = 92 бутылках (группа B).
3. Содержание витаминов в продукте есть случайная величина с функцией распределения F_1 (для группы А) или F_2 (для группы В).
4. Ожидания производителей можно формализовать в виде $F_1(x) = F_2(x)$ для $\forall x > 0$ (т.е. содержание витаминов стохастически одинаково). Т.о., нулевая гипотеза $\mathbf{H}_0$: $F_1 \equiv F_2$ – гипотеза однородности совокупностей (без альтернативы).
5. Уровень значимости $\alpha = 0.025$.
6. Применим критерий однородности хи-квадрат, основанный статистике X^2, равной сумме квадратов разностей частот попадания данных в r = 10 интервалов группировки. Ожидания компании будут подтверждены, если X^2 примет маленькое значение, т.е. критическая область имеет вид $\{X^2 > C\}$.
7. При справедливости нулевой гипотезы функцию распределения статистики X^2 можно приблизить функцией хи-квадрат распределения $\mathbb{Khi}(x \mid r-1) = \mathbf{P}_{\mathrm{H0}}\{X^2 \le x)$ с $r - 1 = 9$ степенями свободы.
8. Критическая константа $C\alpha$ находится как решение неравенства

$$\mathbf{P}_{\mathrm{H0}}\{X^2 \ge C_\alpha) = 1 - \mathbb{Khi}(C\alpha \mid 9) = 0.025,$$

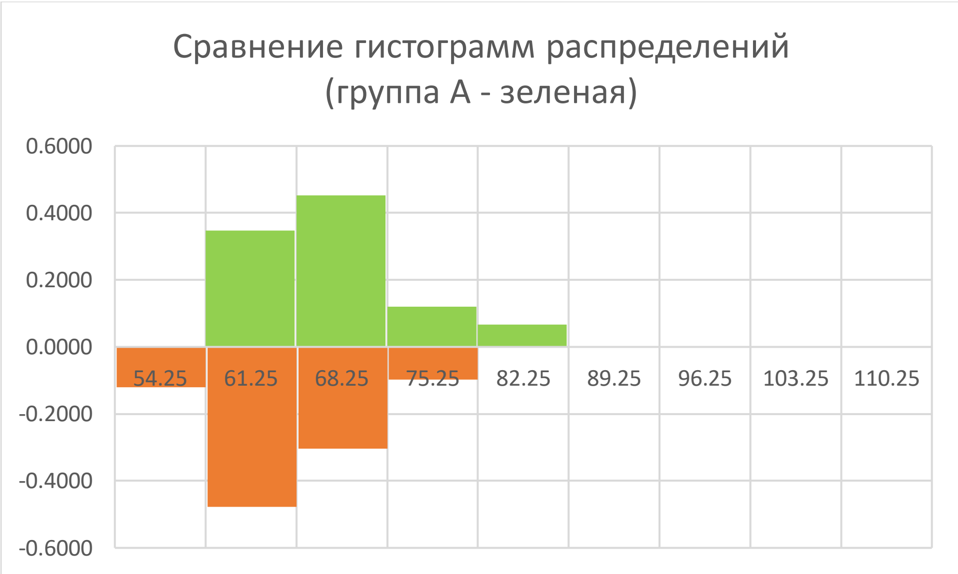
т.е. C_α равна квантили порядка 0.975 хи-квадрат распределения с 9 ст. св. По таблице хи-квадрат распределения (процедурой Excel), находим C_α = 19,02.

   a. Вид критерия: гипотеза однородности отвергается, если $\{X^2 \ge$ 19,02}.

9. По представленным данным:

	Частоты				
Границы	Группа A		Группа B		χ^2
54,25	0	0	11	0,1196	8,967
61,25	26	0,3467	44	0,4783	1,707
68,25	34	0,4533	28	0,3043	2,470
75,25	9	0,1200	9	0,0978	0,188
82,25	5	0,0667	0	0	6,133
89,25	0	0	0	0	0

96,25	0	0	0	0	0
103,25	0	0	0	0	0
110,25	0	0	0	0	0
∞	1	0,0133	0	0	1,227
Σ	75	1	92	1	**20,69**
2,5%-я критическая область					$X^2 \geq 19{,}02$
Вывод	Гипотеза однородности групп				отвергается
p-value	с критическим уровнем значимости				0,0141
Вывод. Содержание витамина С различаются по способу разлива.					



a. p-value вычисляется по формуле p-val = 1 − $\mathbb{Khi}$(20,69| 9) = 0,0141. Есть все основания считать не идентичными способы разлива продукта (П-значение p-val = 0,0141).