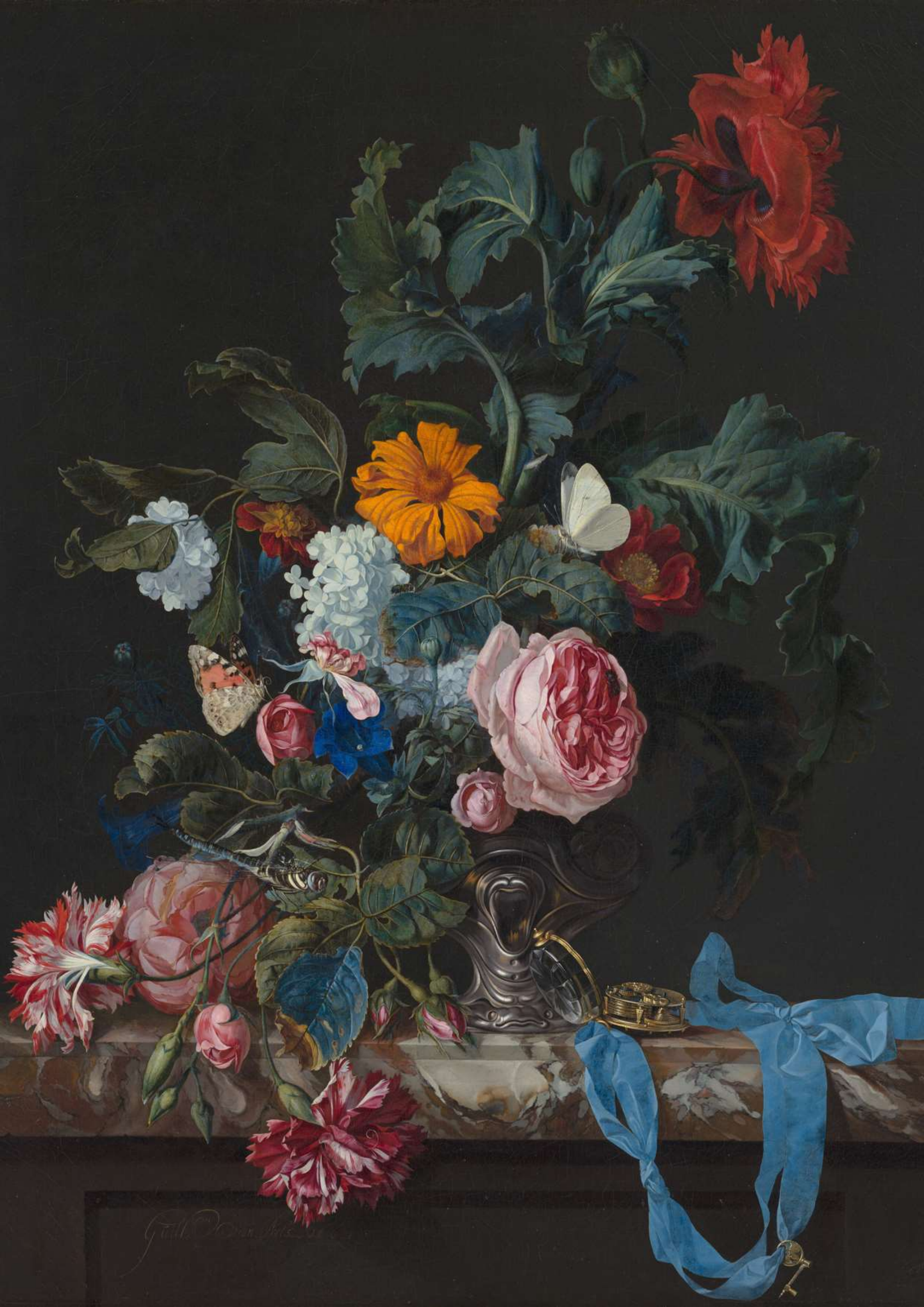

6. При эпидемии гриппа из 200 контролируемых людей однократное заболевание наблюдалось у 181 человека, а дважды болели гриппом 9 человек. Правдоподобна ли гипотеза о том, что в течение эпидемии гриппа число заболеваний отдельного человека суть случайная величина, подчиняющаяся биномиальному распределению с количеством испытаний $n = 2$?

A_0 – не болел A_1 – заболел 1 раз

A_2 – заболел 2 раза

полная группа событий

A – заболел

H_0: $\xi \sim B(2, P(A))$ $\qquad H_1 = \overline{H_0}$

$$1 = C_2^0 P(A)^2 + C_2^1 P(A)(1-P(A)) + C_2^2 (1-P(A))^2$$

$$P(A) = p \qquad P(\overline{A}) = 1-p$$

$$P(A_0) = (1-p)^2 \qquad P(A_1) = 2p(1-p) \qquad P(A_2) = p^2$$

$$\nu = \frac{10}{200} \qquad \nu_1 = \frac{181}{200} \qquad \nu_2 = \frac{9}{200}$$

$$\Delta = n \cdot \sum_{i=0}^{2} \frac{1}{P(A_i)} \left(P(A_i) - \nu_i \right)^2 = n\left(\frac{\left((1-p)^2 - \frac{10}{200}\right)^2}{(1-p)^2} + \frac{\left(2p(1-p) - \frac{181}{200}\right)^2}{2p(1-p)} + \frac{\left(p^2 - \frac{9}{200}\right)^2}{p^2} \right)$$

$$L = \left((1-p)^2\right)^{10} \left(2p(1-p)\right)^{181} \left(p^2\right)^9 \qquad \frac{\delta \ln L}{\delta p} = \frac{-201}{1-p} + \frac{199}{p} = 0 \to \tilde{p} = \frac{199}{400}$$

$$\frac{\delta^2 L}{\delta p^2} = -\frac{201}{(1-p)^2} - \frac{199}{p^2} < 0 \to \text{лок. макс}$$

$$\tilde{\Delta} = n\left(\frac{\left((1-\tilde{p})^2 - \frac{10}{200}\right)^2}{(1-\tilde{p})^2} + \frac{\left(2\tilde{p}(1-\tilde{p}) - \frac{181}{200}\right)^2}{2\tilde{p}(1-\tilde{p})} + \frac{\left(\tilde{p}^2 - \frac{9}{200}\right)^2}{\tilde{p}^2} \right) \simeq 131,23466$$

$$\rightsquigarrow \chi^2(3-1-1) = \chi^2(1)$$

(сложная гипотеза)

$$p\text{-value} = \int_{131,2}^{\infty} \rho_{\chi^2(1)}\,dx = \quad \Bigg| \quad \rho_{\chi^2(1)}(x) = \frac{\left(\frac{1}{2}\right)^{1/2}}{\Gamma(1/2)} x^{\left(\frac{1}{2}-1\right)} \cdot e^{-\frac{1}{2}x} = \frac{e^{-\frac{x}{2}}}{\sqrt{2\pi x}}$$

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} \frac{e^{-x}}{\sqrt{x}}\,dx = 2\int_0^{\infty} e^{-t^2}\,dt = \frac{\sqrt{\pi}}{2}\cdot 2 = \sqrt{\pi} \quad \Bigg|$$

$$= \int_{131,2}^{\infty} \frac{e^{-x/2}}{\sqrt{2\pi x}}\,dx = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{131,2}}^{\infty} \frac{e^{-t^2/2}}{t}\,2t\,dt = \sqrt{\frac{2}{\pi}} \int_{\sqrt{131,2}}^{\infty} e^{-t^2/2}\,dt = \frac{2}{\sqrt{\pi}} \int_{\frac{\sqrt{131,2}}{\sqrt{2}}}^{\infty} e^{-u^2}\,du$$

$(\sqrt{x} = t)$ $\quad \frac{t}{\sqrt{2}} = u$

```
In 32
def destiny(x):
    return exp(-x ** 2)

I = quad(destiny, sqrt(131.2 / 2), float('inf'))
result = sqrt(2 / pi) * I[0]
result
Executed at 2024.03.30 17:00:10 in 6ms

Out 32    1.5830359799410554e-30
```

$< 0{,}05 \rightarrow$ уверенно отвергаем гипотезу

7. Произведено измерение размеров деталей в двух партиях по 100 деталей в каждой партии. В первой партии оказалось 25 деталей с заниженным размером, 50 деталей с точным размером, 25 деталей с завышенным размером, а во второй партии аналогичные числа оказались равны 52, 41, 7 соответственно. Проверить гипотезу о независимости номера партии деталей и размера детали. $: H_0$

ξ – размер детали $\qquad \eta$ – номер партии $\qquad \overline{H_0} = H_1$

$\xi \backslash \eta$	1	2	
корот	25	52	$\frac{77}{200}$
норм	50	41	$91/200$
длин	25	7	$32/200$
	$1/2$	$1/2$	

$$\Delta = \sum \left(\frac{m_{ij} - np_i q_j}{np_i q_j}\right)^2 \sim \chi^2((3-1)(2-1)) = \chi^2(2)$$

$$\Delta = \frac{\left(25 - 200\cdot\frac{1}{2}\cdot\frac{77}{200}\right)^2}{200\cdot\frac{1}{2}\cdot\frac{77}{200}} + \frac{\left(50 - 200\cdot\frac{1}{2}\cdot\frac{91}{200}\right)^2}{200\cdot\frac{1}{2}\cdot\frac{91}{200}} + \frac{\left(25 - 200\cdot\frac{1}{2}\cdot\frac{32}{200}\right)^2}{200\cdot\frac{1}{2}\cdot\frac{32}{200}} +$$

$$\frac{\left(52 - 200\cdot\frac{1}{2}\cdot\frac{77}{200}\right)^2}{200\cdot\frac{1}{2}\cdot\frac{77}{200}} + \frac{\left(41 - 200\cdot\frac{1}{2}\cdot\frac{91}{200}\right)^2}{200\cdot\frac{1}{2}\cdot\frac{91}{200}} + \frac{\left(7 - 200\cdot\frac{1}{2}\cdot\frac{32}{200}\right)^2}{200\cdot\frac{1}{2}\cdot\frac{32}{200}} \approx 20{,}486$$

$$p\text{-value} = \int_{20{,}486}^{\infty} \rho_{\chi^2(2)}\, dx = \quad \Big| \quad \rho_{\chi^2(2)}(x) = \frac{\left(\frac{1}{2}\right)^{2/2}}{\Gamma(2/2)}\, x^{\left(\frac{2}{2}-1\right)} \cdot e^{-\frac{1}{2}x} = \frac{1}{2} e^{-\frac{1}{2}x}$$

$$p\text{-value} = \frac{1}{2}\int_{20{,}486}^{\infty} e^{-\frac{1}{2}x}\, dx =$$

```
In 8   def density(x):
           return 0.5 * exp(-0.5 * x)
       Executed at 2024.03.30 17:30:40 in 14ms

In 9   I = quad(density, 20.483, float('inf'))
       I
       Executed at 2024.03.30 17:30:41 in 17ms

Out 9  (3.5659320517598606e-05, 1.2264749362750056e-09)
```

$< 0.05 \to$ отвергаем уверенно гипотезу

8. Участники олимпиады по математике разбиты на два потока по 300 человек в каждом. Итоги олимпиады оказались следующими: в первом потоке оценки 2, 3, 4, 5 получили соответственно 33, 43, 80 и 144 человека, соответствующие данные для второго потока 39, 35, 72 и 154. Можно ли считать оба потока однородными? H_0

$H_1 = \overline{H_0}$

$n = 600$

	A_2	A_3	A_4	A_5
1 поток	33	43	80	144
2 поток	39	35	72	154
	$\frac{72}{600}$	$\frac{78}{600}$	$\frac{152}{600}$	$\frac{298}{600}$

$$\Delta_1 = \sum \frac{\left(n_{ij} - n_i \hat{P}(A_j)\right)^2}{n_i \hat{P}(A_j)} \sim \chi^2\left((4-1)\cdot(2-1)\right) = \chi^2(3)$$

$$\Delta_1 = \frac{\left(33 - 300\cdot\frac{72}{600}\right)^2}{300\cdot\frac{72}{600}} + \frac{\left(43 - 300\cdot\frac{78}{600}\right)^2}{300\cdot\frac{78}{600}} + \frac{\left(80 - 300\cdot\frac{152}{600}\right)^2}{300\cdot\frac{152}{600}} + \frac{\left(144 - 300\cdot\frac{298}{600}\right)^2}{300\cdot\frac{298}{600}} \approx 1{,}038$$

$$\Delta_2 = \frac{\left(39 - 300\cdot\frac{72}{600}\right)^2}{300\cdot\frac{72}{600}} + \frac{\left(35 - 300\cdot\frac{78}{600}\right)^2}{300\cdot\frac{78}{600}} + \frac{\left(72 - 300\cdot\frac{152}{600}\right)^2}{300\cdot\frac{152}{600}} + \frac{\left(154 - 300\cdot\frac{298}{600}\right)^2}{300\cdot\frac{298}{600}} \approx 1{,}038$$

$$\Gamma\left(\frac{1}{2}+1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$\Delta = \Delta_1 + \Delta_2 \approx 2{,}076$$

$$p\text{-value} = \int_{2{,}076}^{\infty} \rho_{\chi^2(3)}\, dx = \Bigg|$$

$$\rho_{\chi^2(3)}(x) = \frac{\left(\frac{1}{2}\right)^{3/2}}{\Gamma(3/2)} x^{\left(\frac{3}{2}-1\right)} \cdot e^{-\frac{1}{2}x} = \frac{2\sqrt{x}}{2\sqrt{2\pi}} e^{-\frac{1}{2}x} = \frac{\sqrt{x}}{\sqrt{2\pi}} e^{-\frac{1}{2}x}$$

$$p\text{-value} = \int_{2{,}076}^{\infty} \frac{\sqrt{x}}{\sqrt{2\pi}} e^{-\frac{1}{2}x}\, dx$$

```
In 20
def drrkkm_density(x):
    return sqrt(x) / (sqrt(2 * pi)) * exp(-0.5 * x)
Executed at 2024.03.30 18:04:58 in 3ms

In 13
I = quad(drrkkm_density, 2.076, float('inf'))
I
Executed at 2024.03.30 18:00:50 in 13ms

Out 13
(0.5567833651708627, 9.725444059192558e-12)
```

> 0.05 – уверенно отвергаем гипотезу

9. При снятии показаний измерительного прибора десятые доли деления шкалы прибора оцениваются «на глаз» наблюдателем. Количества цифр 0, 1, 2, ..., 9, записанных наблюдателем в качестве десятых долей при 100 независимых измерениях, равны 5, 8, 6, 12, 14, 18, 11, 6, 13, 7 соответственно.

a) Проверить гипотезу о согласии данных с законом равномерного распределения с помощью критерия χ^2 и с помощью критерия Колмогорова. Сравнить результаты.

b) Проверить гипотезу о согласии данных с законом нормального распределения с помощью критерия χ^2 (оценки неизвестных параметров определить численно, максимизируя функцию правдоподобия, построенную по группированной выборке) и с помощью критерия Колмогорова (распределение критерия определить бутстрапом). Сравнить результаты.

а) Критерий Хихи^2

A_i	0	1	2	3	4	5	6	7	8	9
m	5	8	6	12	14	18	11	6	13	7
p_i	$1/10$	$1/10$	$1/10$	$1/10$	$1/10$	$1/10$	$1/10$	$1/10$	$1/10$	$1/10$

$n = 100$

$$\Delta = \sum \frac{(m_i - np_i)^2}{np_i} = 16{,}4 \qquad \tilde{\Delta} \sim \chi^2(9)$$

$$p\text{-value} = \int_{16,4}^{+\infty} q_{\chi^2(9)}\, dx = 0{,}059 > 0{,}05$$ – не отвергаем

Критерий Колмогорова

$p\text{-value} = 0{,}0328 < 0{,}05$ – отвергаем H_0

Колмогоров более мощный и мы отвергли H_0, не смотря на то, что по хи-хи квадрату не отвергли.