Задача 1

- Основная гипотеза H_0: пол абитуриента не влияет на выбор специальности ($p_1 = p_2$)
- Альтернативная гипотеза H_1: пол абитуриента влияет на выбор специальности ($p_1 \neq p_2$)

Таблица сопряженности:

Пол	Точн. науки	Гум. науки	Итого
Муж	197	158	355
Жен	102	221	323
Итого	299	379	678

Используем критерий X^2 для проверки независимости

Вычислим частоты

Пол	Точн. науки	Гуман. науки
М	156,5	198,5
Ж	142,5	180,5

$$М, \text{ Точные} = \frac{355 \times 299}{678} = 156,5$$

$$М, \text{ гуман.} = (355 \times 379)/678 = 198,5$$

$$Ж. \text{ Точн} = (323 \times 299)/678 = 142,5$$

$$Ж. \text{ гуман.} = (323 \times 379)/678 = 180,5$$

$$X^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{(197-156,5)^2}{156,5} + \frac{(158-198,5)^2}{198,5} + \frac{(102-142,5)^2}{142,5} +$$

$$+ \frac{(221-180,5)^2}{180,5} \approx 10,48 + 8,26 + 11,49 + 9,07 = 39,30$$

При верности H_0 статистика X^2 имеет распределение X^2 с $(r-1) \times (c-1)$ степенями свободы, т.е. $1 \times 1 = 1$ степ. своб.

Для уровня значимости $\alpha = 0,05$ и 1 степени свободы критическое значение $X^2 \approx \pm 3,84$

Т.к $39,30 > 3,84$, отвергаем H_0 и принимаем H_1: пол абитуриента влияет на выбор специальности

Коэффициент контингенции:

$$C = \sqrt{\frac{X^2}{X^2 + n}} = \sqrt{\frac{39,30}{39,30 + 678}} \approx 0,236 \Rightarrow \text{зависимость есть}$$

Задача 2

H_0: Нет зависимости между местами от жюри и местами от зрителей $\rho = 0$

H_1: Есть зависимость между местами от жюри и местами от зрителей $\rho \neq 0$

Участница	жюри	Зритель	Разность d	d^2
1	8	10	−2	4
2	2	7	−5	25
3	9	8	1	1
4	6	2	4	16
5	4	5	−1	1
6	5	1	4	16
7	3	6	−3	9
8	7	3	4	16
9	10	9	1	1
10	1	4	−3	9

Сумма квадратов разностей $\sum d_i^2 = 4+25+1+16+1+16+9+16+1+9 = 98$

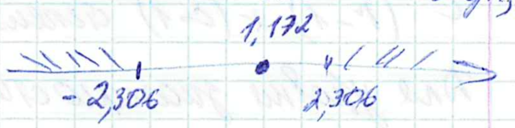
Посчитаем коэф. корел. Спирмена:

$$\hat{\rho}_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6 \times 98}{10(10^2-1)} \approx 1 - 0{,}594 \approx 0{,}406$$

t-статистика для проверки значимости: $t = \hat{\rho}_s\sqrt{\frac{n-2}{1-\hat{\rho}_s^2}}$

$$= 0{,}406 \times \sqrt{\frac{10-2}{1-0{,}406^2}} \approx 0{,}406 \times 2{,}886 \approx 1{,}172$$

Для уровня значимости $\alpha = 0{,}05$ и $n-2 = 10-2 = 8$ степеней свободы, критич. значение $t \approx \pm 2{,}306$



Т.к. $t = 1{,}172$ меньше $2{,}306$, то мы не можем отвергнуть H_0. Т.е. на уровне значимости 0,05 нет оснований считать, что мнения жюри и зрителей зависимы.

Т.к. гипотеза не отвергается, коэф. корел. Спирмана $\hat{\rho}_s \approx 0{,}406$ указывает на умеренную положительную зависимость, но эта зависимость статистически незначима

Задача 3

H_0: нет зависимости между длиной головы (X) и длиной грудного плавника $\rho = 0$

H_1: есть зависимость между длиной головы (X) и длиной грудного плавника $\rho \neq 0$

① Признаки распределены нормально $\Rightarrow$ коэф. корел. Пирсона

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \times \sum (Y_i - \bar{Y})^2}}$$

$$\bar{X} = (66+61+67+73+51+59+48+47+58+44+41+54+52+47+51+45)/16 =$$

$$\bar{X} = 852/16 = 53{,}25$$

$$\bar{Y} = (38+31+36+43+29+33+28+25+36+26+21+30+20+27+28+26)/16 =$$

$$\bar{Y} = 468/16 = 29{,}25$$

X_i	Y_i	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X}) \times (Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$
66	38	12,75	8,75	111,56	162,56	76,56
61	31	7,75	1,75	13,56	60,06	3,06
67	36	13,75	6,75	92,81	189,06	45,56
73	43	19,75	13,75	271,56	390,06	189,06
51	29	-2,25	-0,25	0,56	5,06	0,06
59	33	5,75	3,75	25,56	33,06	14,06
48	28	-5,25	-1,25	6,56	27,56	1,56
47	25	-6,25	-4,25	26,56	39,06	18,06
58	36	4,75	6,75	32,06	22,56	45,56
44	26	-9,25	-3,25	30,06	85,56	10,56
41	21	-12,25	-8,25	101,06	150,06	68,06
54	30	0,75	0,75	0,56	0,56	0,56
52	20	-1,25	-9,25	11,56	1,56	85,56
47	27	-6,25	-2,25	14,06	39,06	5,06
51	28	-2,25	-1,25	2,81	5,06	1,56
45	26	-8,25	-3,25	26,81	68,06	10,56

$$\sum (X_i - \bar{X}) \times (Y_i - \bar{Y}) = 783{,}72$$

$$\sum (X_i - \bar{X})^2 = 1273{,}62$$

$$\sum (Y_i - \bar{Y})^2 = 574{,}36$$

Коэффициент Пирсона

$$r = \frac{783{,}72}{\sqrt{1273{,}62 \times 574{,}36}} \approx 0{,}918$$

$\Rightarrow$ сильная положительная линейная зависимость

Ранг X_i	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
Ранг Y_i	16	14	13	15	12	11	10	8	13	9	5	6	1	7	10	9
d_i Разность ранг.	0	1	1	-2	0	0	0	1	-5	-2	1	-1	3	-4	-8	-8
d_i^2	0	1	1	4	0	0	0	1	25	4	1	1	9	16	64	64

$\sum d_i^2 = 191$ коэф. Спирмена $\hat{\rho}_s = 1 - \frac{6 \times 191}{16(16^2-1)} \approx 1 - 0{,}2813 \approx 0{,}7187$

$\Rightarrow$ сильная положительная монотонная зависимость

② Зависимость ненормальная и немонотонная ⇒ вычисляем коэффициент корреляции Кендала (τ)

Количество согласованных пар (C) = 84, несогласованных (D) = 36

C - порядок значений X_i и Y_i совпадает, D - не совпадает

T_x и T_y - количество пар, с одинаковыми значениями X и Y соот.

$$\tau = \frac{C-D}{\sqrt{(C+D+T_x)\times(C+D+T_y)}} = \frac{84-36}{\sqrt{(84+36)(84+36)}} = \frac{48}{120} = 0,4$$

коэф. корел. Кендала $\tau \approx 0,4$ указывает на умеренную положительную зависимость

③ коэф. корреляции Пирсона:

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0,918 \times \sqrt{\frac{16-2}{1-0,918^2}} = 0,918 \times 4,465 \approx 4,1$$

При $n-2 = 16-2 = 14$ степенях свободы и на уровне значимости $\alpha = 0,05$, критич. значение $t \approx \pm 2,145$

Т.к. $4,1 > 2,145$, H_0 отвергается.

Во всех трех случаях, предположения подтверждаются, существует положительная зависимость между длиной головы и длиной грудного плавника у окуней