Домашнее задание по статистике 03.11.17

1.

$C_0 = 6$, $C_1 = 7$, $C_2 = 6$, $C_3 = 6$, $C_4 = 7$, $C_5 = 6$, $C_6 = 6$, $C_7 = 6$
$k = 8; n = \sum\limits_{i=0}^{7} C_i = 50$
$T = 50 \cdot \sum\limits_{i=0}^{7} \frac{(C_i/50 - 1/8)^2}{1/8} = 0.24$
$T \sim \chi^2(k-1) = \chi^2(7)$
$pvalue = 1 - F_{\chi^2(7)}(T) \approx 0.999953 > 0.05 \Rightarrow$ гипотеза не отвергается при $\alpha = 0.05$

2.

$C_0 = 60, C_1 = 140, C_2 = 125, C_3 = 155$
$k = 4; n = 480; \overline{X} = \frac{960}{480}$
$H_0:$ генеральная совокупность $\xi \sim P(\overline{X}) = P(2)$

$P_0 = P_{H_0}[\xi = 0] = 0.135$
$P_1 = P_{H_0}[\xi = 1] = 0.271$
$P_2 = P_{H_0}[\xi = 2] = 0.271$
$P_3 = P_{H_0}[\xi \geq 3] = 0.323$
$T = 480 \cdot \sum\limits_{i=0}^{3} \frac{(C_i/480 - P_i)^2}{P_i} = 1.31$
$T \sim \chi^2(k-2) = \chi^2(2)$, т.к. в качестве параметра использовалась оценка максимального правдоподобия
$pvalue = 1 - F_{\chi^2(2)}(1.31) = 0.52 > 0.05 \Rightarrow$ гипотеза не отвергается при $\alpha = 0.05$

3.

$\overline{X} = 39217, S_X = 12210, n_X = n = 100$
$\overline{Y} = 43121, S_Y = 17020, n_Y = n = 100$
$H_0 : \mathbb{E}X = \mathbb{E}Y$
$H_1 : \mathbb{E}X \neq \mathbb{E}Y$
$T = \frac{\overline{X} - \overline{Y}}{S_{X-Y}} \sim t(df)$

$$df = \frac{(S_X^2/n_X + S_2^2/n_Y)^2}{(S_X^2/n_X)^2/(n_X-1) + (S_Y^2/n_Y)^2/(n_Y-1)}$$

Поскольку $n_X = n_Y = n, df = \dfrac{(S_X^2 + S_Y^2)^2(n-1)}{S_X^4 + S_Y^2} = \dfrac{99 \cdot (12210^2 + 17020^2)^2}{12210^4 + 17020^4} =$ 179.56

$$S_{X-Y} = \sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}} = \sqrt{\frac{12210^2}{100} + \frac{17020^2}{100}} = 2094$$

$$T = \frac{39217 - 43121}{2094} = -1.864$$

T попадает в 99%-процентный доверительный интервал $= (-2.6; 2.6)$

Таким образом, гипотеза не отвергается

4 (метод генерации нормальных распределений с заданной корреляцией)

$X \sim N(0,1)$, $Z \sim N(0,1)$, $Y = \rho X + \sqrt{1-\rho^2}Z$

$r(X,Y) = \frac{cov(X,Y)}{\sigma X \sigma Y} = cov(X,Y) = cov(X, \rho X) + cov(X, \sqrt{1-\rho^2}Z)$
$= \rho \cdot cov(X,X) = \rho$