Санкт-Петербургский Государственный Университет

Поляков Иван Михайлович

Отчёт по Лабораторной Работе № 5_1

Критерии Стьюдента, Фишера-Снедекора и Колмогорова

Направление 01.04.02: «Прикладная математика и информатика»
Образовательная программа ВМ.5505.2021: «Математическое и информационное обеспечение экономической деятельности»

Преподаватель:
доктор технических наук,
профессор Буре Владимир Мансурович

Санкт-Петербург
2022 г.

1 Постановка задачи

В таблице приводятся сведения об экспорте и импорте Бельгии

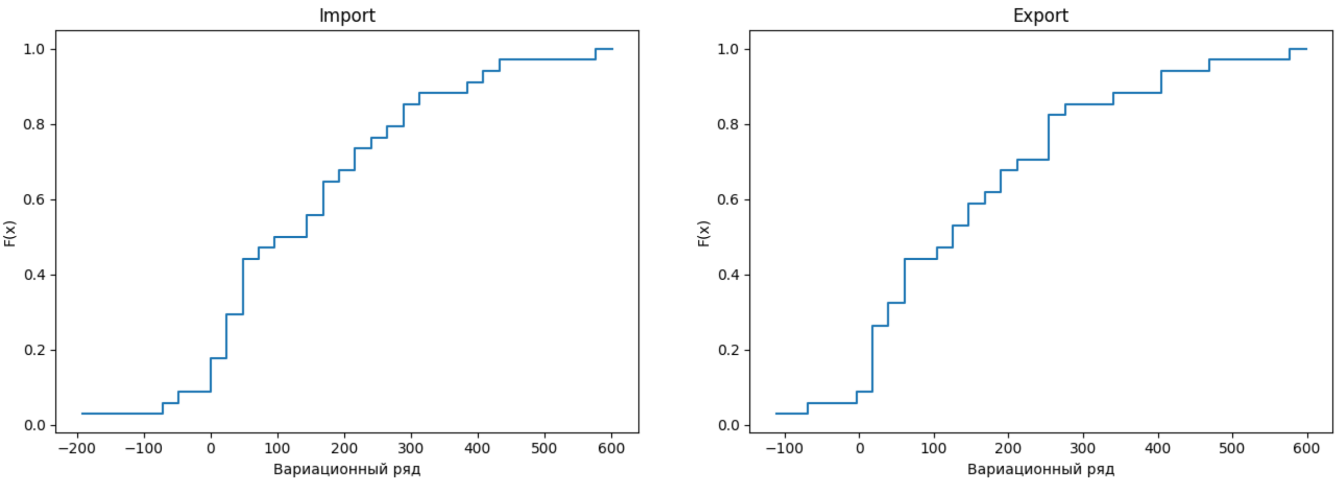
```
Исходные данные
1961: [202.31 209.18]
1962: [219.08 221.57]
1963: [239.85 248.39]
1964: [278.27 283.73]
1965: [306.39 305.59]
1966: [328.61 337.14]
1967: [352.37 351.59]
1968: [402.38 400.18]
1969: [483.37 474.39]
1970: [562.65 533.68]
1971: [609.19 581.12]
1972: [683.46 633.35]
1973: [846.47 811.16]
1974: [1116.27 1109.35]
1975: [1065.21 1061.39]
1976: [1266.58 1261.47]
1977: [1474.85 1499.88]
1978: [1540.11 1570.85]
1979: [1798.81 1866.38]
1980: [2026.65 2125.11]
1981: [2286.64 2357.69]
1982: [2640.57 2694.44]
1983: [2924.28 2864.72]
1984: [3337.84 3277.94]
1985: [3479.21 3379.87]
1986: [3367.92 3187.85]
1987: [3477.43 3334.15]
1988: [3900.28 3719.21]
1989: [4498.77 4320.59]
1990: [4660.38 4506.95]
1991: [4846.69 4658.41]
1992: [4980.87 4713.19]
1993: [5012.75 4674.73]
1994: [5491.28 5108.09]
1995: [5764.45 5377.49]
```

Считая, что выборки первых разностей подчиняются нормальному распределению, на уровне значимости 5% необходимо проверить равенство средних и дисперсий по критериям Стьюдента и Фишера-Снедекора. Необходимо проверить нормальность выборок по критерию Колмагорова для сложных гипотез.

2 Ход работы

Изначально были изучены разделы 2.4.1, 2.4.2, 2.4.5, 2.4.8 из книги Буре В.М., Парилина Е.М., Седаков А.А. «Методы прикладной статистики в R и Excel».

Была произведена обработка данных: вычислены первые разности и исключены повторения. На основе полученных временных рядов получены следующие функции распределения:



Дальнейшие вычисления по критериям производились в соответствии с алгоритмами, описывающие критерии.

2.1 Критерий Стьюдента

Алгоритм критерия Стьюдента в предположении известных дисперсий

1. Выдвинуть гипотезу $H_0 : a_1 = a_2$. Сформулировать альтернативную гипотезу $H_0 : a_1 \neq a_2$
2. Задать уровень значимости критерия α
3. Вычислить значение статистики по формуле

$$\phi_1\left(X_{[n]}, Y_{[m]}\right) = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}, \tag{1}$$

   которая подчиняется стандартному нормальному распределению, при условии, что справедлива гипотеза H_0.

4. Найти критическую область

$$\left(-\infty; -u_{1-\frac{\alpha}{2}}\right) \cup \left(u_{1-\frac{\alpha}{2}}; \infty\right), \tag{2}$$

   где $u_{1-\frac{\alpha}{2}}$ – квантиль стандартного нормального распределения уровня $1 - \frac{\alpha}{2}$

5. Если модуль численного значения статистики (1) попадает в интервал $\left(u_{1-\frac{\alpha}{2}}; \infty\right)$, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований её отвергнуть при уровне значимости α.

Алгоритм критерия Стьюдента в предположениях неизвестных, но равных дисперсиях

1. Выдвинуть гипотезу $H_0 : a_1 = a_2$. Сформулировать альтернативную гипотезу $H_0 : a_1 \neq a_2$
2. Задать уровень значимости критерия α
3. Вычислить значение статистики по формуле

$$\phi_2\left(X_{[n]}, Y_{[m]}\right) = \frac{\overline{x} - \overline{y}}{\hat{s}\sqrt{\frac{1}{n} + \frac{1}{m}}}, \tag{3}$$

   где

$$\hat{s} = \sqrt{\frac{ns_1^2 + ms_2^2}{n+m-2}} = \sqrt{\frac{(n-1)\tilde{s}_1^2 + (m-1)\tilde{s}_2^2}{n+m-2}}, \tag{4}$$

где s_1^2, s_2^2 – выборочные дисперсии, $\tilde{s}_1^2$, $\tilde{s}_2^2$ – исправленные выборочные дисперсии выборок $X_{[n]}, Y_{[m]}$

4. Найти критическую область

$$\left(-\infty; -t_{1-\frac{\alpha}{2}, n+m-2}\right) \cup \left(t_{1-\frac{\alpha}{2}, n+m-2}; \infty\right), \tag{5}$$

где $t_{1-\frac{\alpha}{2}, n+m-2}$ – квантиль распределения Стьюдента с $(n+m-2)$ степенями свободы уровня $1-\frac{\alpha}{2}$

5. Если модуль численного значения статистики (3) попадает в интервал $\left(t_{1-\frac{\alpha}{2}, n+m-2}; \infty\right)$, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований её отвергнуть при уровне значимости α.

2.2 Критерий Фишера-Снедекора

Алгоритм критерия Фишера-Снедекора

1. Выдвинуть нулевую гипотезу $H_0 : \sigma_1^2 = \sigma_2^2$. Сформулировать альтернативную гипотезу $H_1 : \sigma_1^2 \neq \sigma_2^2$

2. Задать уровень значимости критерия α

3. Вычислить значение статистики по формуле

$$F(X_{[n]}, Y_{[m]}) = \frac{\frac{s_1^2 n}{n-1}}{\frac{s_2^2 m}{m-1}} = \frac{\tilde{s}_1^2}{\tilde{s}_2^2}, \tag{6}$$

где s_1^2, s_2^2 – выборочные дисперсии выборок $X_{[n]}, Y_{[m]}$

4. Найти критическую область

$$\left[0; F_{\frac{\alpha}{2}, n-1, m-1}\right) \cup \left(F_{1-\frac{\alpha}{2}, n-1, m-1}; \infty\right], \tag{7}$$

где $F_{\frac{\alpha}{2}, n-1, m-1}$ и $F_{1-\frac{\alpha}{2}, n-1, m-1}$ – квантили уровней $\frac{alpha}{2}$ и $1-\frac{alpha}{2}$ распределения Фишера с $(n-1)(m-1)$ степенями свободы

5. Если значение статистики (6) попадает в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет основания её отвергнуть при уровне значимости α.

2.3 Критерий Колмогорова

Алгоритм критерия солгасия Колмогорова в случае сложной гипотезы о нормальности распределения генеральной совокупности

1. Выдвинуть нулевую гипотезу $H_0 : F_\xi(\cdot) = F_0(\cdot, \theta)$. Сформулировать альтернативную гипотезу $H_1 : F_\xi(\cdot) \neq F_0(\cdot, \theta)$;

2. Задать уровень значимости критерия α;

3. Найти оценки $\hat{\theta} = (\overline{x}, s^2)$ неизвестных параметров распределения $\theta = (a, \sigma^2)$;

4. Вычислить значение исправленной формы статистики $\tilde{D}_n^*$ следующим образом:

   (a) По выборке $X_{[n]}$ построить эмпирическую функцию распределения $F_n(x)$ по формуле:

$$F_n(x) = \begin{cases} 0, & x < x_1, \\ \frac{1}{n}, & x_1 \le x < x_2, \\ \vdots \\ \frac{k}{n}, & x_k \le x < x_{k+1}, \\ \vdots \\ 1, & x \ge x_n \end{cases} \tag{8}$$

   (b) Определить D_n^* по формуле

$$D_n^* = \max_{1 \le i \le n} \left\{ \frac{i}{n} - F_0(x_i), F_0(x_i) - \frac{i-1}{n} \right\} \tag{9}$$

   (c) Вычислить значение исправленной формы модифицированной статистики Колмогорова по формуле

$$\tilde{D}_n^* = D_n^* \left(\sqrt{n} - 0.01 + \frac{0.85}{\sqrt{n}} \right) \tag{10}$$

5. Найти критическую область - интервал $(d_{1-\alpha}; \infty)$. Квантиль $d_{1-\alpha}$ можно найти из таблицы ниже

Модифицированная форма	0.15	0.1	0.05	0.025	0.01
$D_n^*(\sqrt{n} - 0.01 + \frac{0.85}{\sqrt{n}})$	0.775	0.819	0.895	0.955	1.035

6. Если численное значение статистики $\tilde{D}_n^*$ попадает в интервал $(d_{1-\alpha}; \infty)$, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований отвергнуть нулевую гипотезу при уровне значимости приближённо равном α.

Результаты работы всех критериев при уровнях значимости $\alpha = 0.05$ представлены ниже.

```
Критерий Стьюдента, известные дисперсии
Гипотеза H0 принимается на уровне значимости 0.05,
так как значение phi1 = |0.3005| не принадлежит интервалу (1.96; +inf)

Критерий Стьюдента, неизвестные дисперсии
Гипотеза H0 принимается на уровне значимости 0.05,
так как значение phi2 = |0.2961| не принадлежит интервалу (1.96; +inf)

Критерий Фишера-Снедекора, неизвестные дисперсии
Гипотеза H0 принимается на уровне значимости 0.05,
так как значение F = 0.9472 не принадлежит интервалу [0; 0.4994) U (2.0023; +inf)

Экспорт
Критерий Колмогорова
Гипотеза H0 отвергается на уровне значимости 0.05,
так как значение 1.75492 попадает в интервал (0.895; +inf)

Импорт
Критерий Колмогорова
Гипотеза H0 отвергается на уровне значимости 0.05,
так как значение 1.40394 попадает в интервал (0.895; +inf)
```