Занятие седьмое. О мерах зависимости и корреляционных расстояних, о корреляции около кривой и главных компонентах.

В рамках этого занятия мы будем заниматься вопросами зависимости и независимости. Простейшей мерой выражения зависимости для нас является коэффициент корреляции

$$\rho_{X,Y} = \frac{\mathbf{E}(X - \mathbf{E}X)(Y - \mathbf{E}Y)}{\sqrt{\mathbf{D}X\mathbf{D}Y}}.$$

Этот коэффициент $\rho \in [-1, 1]$, причем для независимых величин $\rho = 0$, а для линейно зависимых величин (и только для них) $\rho = 1$ или $\rho = -1$. Коэффициент ρ хорошо отражает прямую (линейную зависимость), но не отслеживает более сложной зависимости.

Вопрос 1. Показать, что для $X \sim \mathcal{N}$, $\rho_{X,X^2} = 0$, хотя величины зависимы.

Тем не менее, отличие коэффициента корреляции от нуля показывает нам наличие зависимости. В связи с этим используют оценку для ρ, имеющую вид

$$\hat{\rho} = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\sqrt{(\overline{X^2} - \overline{X}^2)(\overline{Y^2} - \overline{Y}^2)}}.$$

Мы обсудим эту меру зависимости и ее модификации.

Независимость дискретных выборок

Для проверки независимости двух дискретных выборок данные удобно представить в виде так называемых таблиц сопряженности (contingency table). Если X принимает значения $x_1, ..., x_k$, а Y — $y_1, ..., y_l$, то составим таблицу со столбцами $y_1, ..., y_l$ и строками $x_1, ..., x_k$, где в ячейке (i, j) стоит $p_{i,j} = \mathbf{P}(X = x_i, Y = y_j)$. Если гипотеза верна, то $p_{i,j} = p_i q_j$, где $p_i = \mathbf{P}(X = x_i)$, $q_j = \mathbf{P}(Y = y_j)$.

1) Используем параметрический критерий хи-квадрат на основе частот $\nu_{i,j}$ появления пар (x_i, y_j). Тогда гипотеза заключается в том, что наше $kl - 1$ параметрическое семейство вероятностей $p_{i,j}$ выражается через $k + l - 2$ вероятности p_i, q_j, $i < k$, $j < l$. Функция правдоподобия при выполнении гипотезы имеет вид

$$\prod_{i=1}^{k}\prod_{j=1}^{l}(p_i q_j)^{\nu_{i,j}} = \prod_{i=1}^{k} e^{\ln p_i \sum_{j=1}^{l}\nu_{i,j}} \prod_{j=1}^{l} e^{\ln q_j \sum_{i=1}^{k}\nu_{i,j}}$$

Отсюда ОМП $\hat{p}_i = \nu_{i,\cdot}/N = \sum_{j=1}^{l}\nu_{i,j}/N$, $\hat{q}_j = \sum_{i=1}^{k}\nu_{i,j}/N = \nu_{\cdot,j}/N$.

Следовательно, при выполнении гипотезы

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(\nu_{i,j} - N\hat{p}_i\hat{q}_j)^2}{N\hat{p}_i\hat{q}_j} \sim \chi^2_{(k-1)(l-1)}$$

Отсюда возникает критерий хи-квадрат для проверки независимости, который в R задан все тем же chisq.test.

2) Что даст для нашего случая критерий отношения правдоподобий?

$$2\ln\left(\frac{\sup_{p_{i,j}} L(x,p)}{\sup_{p_i,q_j} L(x,p)}\right) = 2\sum_{i,j}\ln\left(\frac{\nu_{i,j}^{\nu_{i,j}}}{(\nu_{i,\cdot}\nu_{\cdot,j}/N)^{\nu_{i,j}}}\right) = 2\sum_{i,j}\nu_{i,j}\ln\left(\frac{\nu_{i,j}N}{\nu_{i,\cdot}\nu_{\cdot,j}}\right) \sim \chi^2_{(k-1)(l-1)}.$$

Тем самым, статистики критериев хи-квадрат и отношения правдоподобий для проверки однородности и независимости одинаковы. Это следствие того, что независимость X и Y равносильна совпадению условных распределений $\mathbf{P}(X = x|Y = y_1)$, $\mathbf{P}(X = x|Y = y_2)$,...,$\mathbf{P}(X = x|Y = y_k)$.

В случае двух выборок одинаковы и рассматриваемые предельные распределения.

Упражнение 1. Исследовать критерии на данных survey пакета MASS. Взаимосвязаны ли курение и выполнение домашних задач? Пол и выполнение д.з.?

Те же методы работают для недискретных распределений за счет группировки данных.

С критериями 1) и 2) связаны коэффициенты, оценивающие взаимосвязь величин X, Y:

1) Коэффициент Крамера (V)

$$V = \frac{\chi^2}{N \min(k-1, l-1)}$$

Этот коэффициент лежит в диапазоне $[0, 1]$, причем 0 получается только при $\frac{n_{i,j}}{N} = \frac{n_{i,\cdot}}{N}\frac{n_{\cdot,j}}{N}$, а 1 только в том случае, когда в каждой строке или столбце таблицы сопряженности не более 1 непустого значения.
2) Коэффициент неопределенности или U-коэффициент Тэйла

$$U_{Y|X} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{l} \nu_{i,j} \ln\left(\frac{\nu_{i,j}N}{\nu_{i,\cdot}\nu_{\cdot,j}}\right)}{\sum_{i=1}^{k} \nu_{i,\cdot} \ln\left(\frac{\nu_{i,\cdot}}{N}\right)}.$$

В силу его асимметричности, зачастую рассматривают коэффицент

$$U_{Y,X} = \frac{U_{Y|X}H_X + U_{X|Y}H_Y}{H_X + H_Y},$$

где $H_X = \sum_{i=1}^{k} \nu_{i,\cdot} \ln\left(\frac{\nu_{i,\cdot}}{N}\right)$, $H_Y = \sum_{i=1}^{k} \nu_{\cdot,j} \ln\left(\frac{\nu_{\cdot,j}}{N}\right)$.
Этот коэффициент также находится в диапазоне $[0, 1]$, причем принимает значения 0 и 1 в тех же случаях, что и V.

Меры связи

Мерами связи (measure of association) называют коэффициенты, аналогичные коэффициентам Крамера и Тэйла, показывающие взаимосвязь. Рассмотрим еще несколько таких мер в общем, недискретном случае.
1) Коэффициент корреляции Пирсона ρ.
Оценим коэффициент корреляции выборочной величиной

$$\rho = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\sqrt{S_X^2 S_Y^2}},$$

где S_X^2, S_Y^2 — смещенные оценки дисперсий $\overline{(X-\overline{X})^2}$, $\overline{(Y-\overline{Y})^2}$. Этот коэффициент лежит в диапазоне $[-1, 1]$, причем 1 будет достигаться только при $X = aY + b$ п.н. при некотором $a > 0$, -1 при $X = aY + b$ п.н., $a < 0$. Теоретический коэффициент корреляции измеряет наличие прямой линейной зависимости, поэтому коэффициент Пирсона также будет принимать большие абсолютные значения именно при линейной зависимости, но хуже будет отслеживать другие виды зависимостей.
Коэффициент Пирсона обычно применяют для двумерных нормальных выборок, для которых он является неплохой мерой зависимости.
2) Коэффициент Спирмена ρ_S.
Этот коэффициент позволяет уйти от условия линейности зависимости, заменив наблюдения X_i на их ранги R_i в ряду X, а Y_i — на их ранги T_i в ряду Y. Тогда

$$\rho_S = \frac{\overline{RT} - \overline{R}\,\overline{T}}{\sqrt{S_R^2 S_T^2}}.$$

Тогда близость ρ_S по абсолютному значению к 1 означает, что R_i линейно зависят от T_i, т.е. зависимость Y от X монотонна.
3) Коэффициент Кенделла τ.
Назовем две пары значений x_i, y_i, x_j, y_j согласованными, если $x_i - x_j$, $y_i - y_j$ — одного знака. Пусть C — количество согласованных, D — несогласованных пар. Тогда

$$\tau = \frac{C - D}{C + D} = \frac{2(C - D)}{n(n-1)}$$

называют коэффициентом согласия Кенделла или τ-коэффициентом. Он также лежит в диапазоне $[-1, 1]$. Этот коэффициент сильно коррелирован с коэффициентом ρ_S.
В R один из способов получения коэффициентов 1)-3) является cor(X,Y,method = c("pearson", "kendall",

"spearman")). Методы CramerV, KendallTauB, SpearmanRho, Lambda, UncertCoef пакета DescTools строят коэффициенты вместе с доверительными интервалами.

Для двумерных нормальных данных удобно использовать так называемое преобразование Фишера $arctanh(z) = 0.5\ln((1+z)/(1-z))$, приводящее коэффициент ρ к асимптотической нормальности со среднем $arctanh(\rho)$ и дисперсией $1/(n-3)$, откуда можно построить доверительный интервал для ρ. Аналогичный результат верен и для коэффициентов Спирмена и Кенделла.

При выполнении гипотезы независимости справедливы сходимости

$$\frac{\rho_S}{\sqrt{\mathbf{D}\rho_S}} = \rho_S\sqrt{n-1} \xrightarrow{d} Z \sim \mathcal{N}(0,1),\ \frac{\tau}{\sqrt{\mathbf{D}\tau}} = \tau\sqrt{\frac{9n(n-1)}{2(2n+5)}} \xrightarrow{d} Z \sim \mathcal{N}(0,1),$$

откуда вытекают асимптотические критерии для проверки гипотезы независимости.
Вопрос 2. Показать, что для наблюдений без повторений $\mathbf{D}\rho_S = 1/(n-1)$.

Эти критерии, а также их точные версии для небольших n реализованы в cor.test пакета stats, где, опять же, можно выбрать конкретный метод. Более точную версию критерия Спирмена можно найти в spearman.test.
Упражнение 2. Исследовать коэффициенты на выборках $X, f(X)+U$, где X, U — нез-ы $\mathcal{N}$ или R (и а) $f(x) = 2x+5$, б) $f(x) = x^2$, в) $f(x) = \ln x$, г) $f(x) = \sin x$.
Упражнение 3. Исследовать коэффициенты на (X,Y), распределенных на кольце $\sqrt{X^2+Y^2} \in [a,b]$, а) $a = 0.5$, $b = 1$, б) $a = 0.75$, $b = 1$.

Для многих выборок ($k \geq 3$) аналогичную роль может сыграть коэффициент конкордации, подсчитанный по многим выборкам. Одним из таких коэффициентов является коэффициент Кенделла

$$W = \frac{12}{k^2(n^3-n)}\sum_{i=1}^{n}\left(\sum_{j=1}^{k} R_{i,j} - \frac{k(n+1)}{2}\right)^2,$$

где $R_{i,j}$ — ранг i-ого элемента j-ой выборки внутри своей выборки. Эта величина лежит в диапазоне $[0,1]$. При больших n $k(n-1)W$ близок по распределению к χ^2_{n-1}. В R этот коэффициент задается функцией KendallW из DescTools, позволяющей также найти соответствующее p-value гипотезы о совместной независимости.

Еще одним важным коэффициентом является частный или очищенный коэффициент корреляции, позволяющий исключить зависимость двух переменных через третью

$$\frac{\rho_{X,Y} - \rho_{X,Z}\rho_{Y,Z}}{\sqrt{(1-\rho^2_{X,Z})(1-\rho^2_{Y,Z})}},$$

который оценивается исходя из выборки с помощью статистики

$$\frac{\hat\rho_{X,Y} - \hat\rho_{X,Z}\hat\rho_{Y,Z}}{\sqrt{(1-\hat\rho^2_{X,Z})(1-\hat\rho^2_{Y,Z})}},$$

где $\hat\rho$ — выборочные коэффициенты корреляции Пирсона. Этот коэффициент эффективно работает для многомерных нормальных данных, для общих данных можно использовать выборочные коэффициенты корреляции Кенделла. В R такие величины могут быть рассчитаны с помощью функции pcor пакета ppcor. Полученная матрица estimate будет содержать исключенные корреляции по каждой паре переменных при условии всех остальных.
Упражнение 4. Исследовать выборки X, Y, Z с помощью частных корреляций: а) $U \sim \mathcal{N}(0,1)$, $V \sim \mathcal{N}(0,1)$, $Z \sim \mathcal{N}(1,1)$, где $X = U+Z$, $Y = V+Z$, б) $U \sim R[0,1]$, $V \sim R[0,1]$, $Z \sim \exp(\lambda)$, $X = U^2Z$, $Y = V^2Z$.

Ковариационные расстояния

Работа с ковариацией затрудняется тем, что она измеряет именно линейную зависимость. Для измерения

общей зависимости можно использовать другие функционалы — ковариационные расстояния.
1) Метод Хёфдинга (Hoeffding).
Этот метод основан на том, что величина

$$D = \int_{\mathbb{R}^2} (F_{X,Y}(x,y) - F_X(x)F_Y(y))^2 dF_{X,Y}(x,y)$$

неотрицательна и равна нулю в абсолютно-непрерывном случае тогда и только тогда, когда величины независимы.
Вопрос 3. Показать, что в дискретном случае это не так.
При этом $D \leq 1/30$. Для оценки D используется довольно сложная оценка, нахождение которое реализовано в R в виде функции hoeffd пакета Hmisc, которая по матрице данных строит матрицу расстояний. Метод Хёфдинга приведен скорее из исторических соображений, нижеследующий метод, по всей видимости, является более современным.
2) Метод Шекели-Риззо (Szekely-Rizzo, 2009)
Этот метод предлагает рассматривать в качестве меры зависимости между случайными векторами $X \in \mathbb{R}^k$, $Y \in \mathbb{R}^l$, имеющими конечное математическое ожидание, величину

$$V^2_{X,Y} = \int_{t\in\mathbb{R}^k, s\in\mathbb{R}^l} |\psi_{X,Y}(t,s) - \psi_X(t)\psi_Y(s)|^2 w(t,s)dtds,$$

обнуляющуюся только при независимых векторах X, Y, где ψ — характеристические функции, w — некоторая (интегрируемая) весовая функция. Тогда

$$\rho_V = \frac{V^2_{X,Y}}{\sqrt{V_{X,X}V_{Y,Y}}}$$

будет принимать значения в отрезке $[0,1]$. В роли $w(t,s)$ предлагается рассматривать

$$\frac{1}{c_k c_l ||t||_k^{1+k} ||s||_l^{1+l}},\ c_m = \frac{\pi^{1+m}}{\Gamma((1+m)/2)},$$

где норма рассматривается Евклидова. Для оценки ρ_V на выборке из n выборок (X,Y) введем

$$a_{i,j} = ||X_i - X_j||_k,\ a_{i,\cdot} = \frac{1}{n}\sum_{j=1}^n a_{i,j},\ a_{\cdot,\cdot} = \frac{1}{n^2}\sum_{i,j} a_{i,j},\ b_{i,j} = ||Y_i - Y_j||_l,\ b_{\cdot,j} = \frac{1}{n}\sum_{i=1}^n b_{i,j},\ b_{\cdot,\cdot} = \frac{1}{n^2}\sum_{i,j} b_{i,j},$$

положим $A_{i,j} = a_{i,j} - a_{i,\cdot} - a_{\cdot,j} + a_{\cdot,\cdot}$, $B_{i,j} = b_{i,j} - b_{i,\cdot} - b_{\cdot,j} + b_{\cdot,\cdot}$. Тогда

$$\hat{\rho}_V = \frac{\sum_{i,j} A_{i,j}B_{i,j}}{\sqrt{\sum_{i,j} A^2_{i,j} \sum_{i,j} B^2_{i,j}}}.$$

Это сильно состоятельная оценка ρ_V.
При этом можно построить асимптотический критерий, основанный на том, что при выполнении гипотезы независимости

$$\limsup_{n\to\infty} \mathbf{P}\left(\frac{n\hat{\rho}_V}{S_2} > z^2_{1-\alpha/2}\right) \leq \alpha,$$

при некотором S_2, где z — квантиль $\mathcal{N}(0,1)$.
В R этот метод реализован функцией dcor, а соответствующий критерий — функцией dcor.ttest пакета Energy.
Упражнение 5. Исследовать метод на тех же примерах, что и в предыдущей части.
В случае многих переменных строят матрицу попарных расстояний и таким образом исследуют, какие переменные сильно зависимы.

Корреляция около кривой

Тема нужна только при сдаче на сложном уровне

Как мы уже говорили, коэффициент корреляции хорошо измеряет наблюдения, распределенные вдоль некоторой прямой. Введем аналогичные понятия, связанные с распределением вдоль кривой. Вектор (X,Y) назовем распределенным вдоль параметрической кривой $(x,y)=c(s)$, где s — натуральный параметр ($|c'(s)|=1$), если $(X,Y)=c(S)+Td(S)$ для некоторого вектора (S,T), где $d(s)$ — единичный нормальный вектор к кривой в точке $c(s)$. При этом будем требовать: а) $\mathbf{E}(T|S)=0$ п.н. б) $\mathbf{D}(T|S)\le \mathbf{D}S$ п.н.

Вопрос 4. Найти распределения S, $T|S$ для выборки

а) равномерно распределенной на квадрате с вершинами $(1,1)$, $(1,-1)$, $(-1,-1)$, $(-1,1)$ для $c(t)=(t,0)$.

б) равномерно распределенной на квадрате с вершинами $(1,0),(0,1),(-1,0),(0,-1)$ для $c(t)=(t,0)$ и для $c(t)=(t,t)/\sqrt{2}$.

в) (S,T) для выборки, равномерно распределенной на кольце $\sqrt{X^2+Y^2}\in[0.5,1]$ для $c(t)=R(\cos at,\sin at)$. Какими должны быть R, a?

Положим $\alpha(t)$ — угол между касательной к $c(t)$ и осью абцисс и введем локальные дисперсии и ковариацию:

$$DLoc_X(s)=\mathbf{D}S\cos^2\alpha(s)+\mathbf{D}(T|S=s)\sin^2\alpha(s),\ DLoc_Y(s)=\mathbf{D}S\sin^2\alpha(s)+\mathbf{D}(T|S=s)\cos^2\alpha(s), \quad (1)$$

$$covLoc_{(X,Y)}(s)=(\mathbf{D}S-\mathbf{D}(T|S=s))\cos\alpha(s)\sin\alpha(s),\ corrLoc_{(X,Y)}(s)=\frac{covLoc_{(X,Y)}}{\sqrt{DLoc_X DLoc_Y}}. \quad (2)$$

В частности, если $c(t)$ — прямая, а $D(T|S=s)$ не зависит от s, то $DLoc_X=DX$, $DLoc_Y=DY$, $covLoc_{X,Y}=cov(X,Y)$.

Ковариацией и корреляцией около кривой называют

$$covGC(X,Y)=\sqrt{\mathbf{E}(covLoc_{X,Y}(S))^2},\ corrGC(X,Y)=\sqrt{\mathbf{E}(corrLoc_{X,Y}(S))^2}.$$

Вопрос 5. Найти ковариацию (X,Y) для равномерного распределения на круге радиуса $3/2$ с $c(t)=(\cos t,\sin t)$.

Для независимых величин эти показатели оказываются нулевыми, зато даже для тех видов зависимости, для которых ковариация нулевая (например, для примера в)), мы можем ее отследить. Алгоритм построения кривой c(t) называют principle curve fitting. В некотором смысле этот подход близок к методу главных компонент, но здесь выделяется не главная прямая, а главная кривая. Программа, вычисляющая covGC и corrGC, находится в приложенном файле CovrGC.

Упражнение 6. Исследовать метод на (X,Y), распределенных на кольце $\sqrt{X^2+Y^2}\in[a,b]$, а) $a=0.5$, $b=1$, б) $a=0.75$, $b=1$.

Ответы на вопросы

1. Отметим, что коэффициент корреляции не меняется при линейных заменах, то есть $corr(aX+b,Y)=corr(X,Y)$. Следовательно, можно считать, что $X\sim\mathcal{N}(0,1)$. Отсюда $cov(X,X^2)=\mathbf{E}X^3-\mathbf{E}X\mathbf{E}X^2=0$, поскольку $\mathbf{E}X=0$, $\mathbf{E}X^3=0$ из симметрии.

2. Заметим, что раз совпадений нет, то R_i принимают значения от 1 до n и S_i тоже. Тогда

$$\rho_S=\frac{\frac{1}{n}\sum_{i=1}^n R_iS_i-\left(\frac{n+1}{2}\right)^2}{\frac{1}{n}\sum_{i=1}^n i^2-\left(\frac{n+1}{2}\right)^2}.$$

При этом,

$$\mathbf{E}R_i^2=\frac{1}{n}\sum_{i=1}^n i^2=\frac{n(n+1)(2n+1)}{6n},\ \mathbf{D}R_i=\frac{n+1}{2}\left(\frac{2n+1}{3}-\frac{n+1}{2}\right)=\frac{(n+1)(n-1)}{12}.$$

Аналогично при $i\ne j$

$$\mathbf{E}R_iR_j=\sum_{i\ne j}\frac{ij}{n(n-1)}=\frac{(\sum_{i=1}^n i)^2-\sum_{i=1}^n i^2}{n(n-1)}=\frac{\frac{n^2(n+1)^2}{4}-\frac{n(n+1)(2n+1)}{6}}{n(n-1)}=\frac{(n+1)(3n+2)(n-1)}{12(n-1)}$$

Отсюда при $i \neq j$

$$cov(R_iS_i, R_jS_j) = (\mathbf{E}R_iR_j)^2 - (\mathbf{E}R_i)^4 = \left(\frac{(n+1)(3n+2)}{12}\right)^2 - \frac{(n+1)^4}{2^4} = \frac{(n+1)^2((3n+2)^2 - 9(n+1)^2)}{12^2},$$

при $i = j$

$$cov(R_iS_i, R_iS_i) = (\mathbf{E}R_i^2)^2 - (\mathbf{E}R_i)^4 = \frac{(n+1)^2(2n+1)^2}{6^2} - \frac{(n+1)^4}{2^4} = \frac{(n+1)^2((4n+2)^2 - (3n+3)^2)}{12^2}.$$

Следовательно,

$$\mathbf{D}\rho_S = \frac{\frac{1}{n^2}\mathbf{D}(\sum_{i=1}^n R_iS_i)}{\frac{(n+1)^2(n-1)^2}{12^2}} = \frac{\sum_{i,j=1}^n cov(R_iS_i, R_jS_j)}{n^2(n+1)^2(n-1)^2/12^2} = \frac{-n(n-1)\frac{(n+1)^2(6n+5)}{12^2} + n\frac{(n+1)^2(n-1)(7n+5)}{12^2}}{n^2(n+1)^2(n-1)^2/12^2} = \frac{1}{n-1}.$$

3. Рассмотрим вектор (X, Y), принимающий значения $(0, 1)$ и $(1, 0)$. Тогда из условия

$$\int_{\mathbb{R}^2} |F_{X,Y}(x,y) - F_X(x)F_Y(y)| dF_{X,Y}(x,y) = |F_{X,Y}(1,0) - F_X(1)F_Y(0)| + |F_{X,Y}(0,1) - F_X(0)F_Y(1)| =$$
$$|\mathbf{P}((X,Y) = (1,0)) - \mathbf{P}((X,Y) = (1,0))| + |\mathbf{P}((X,Y) = (0,1)) - \mathbf{P}((X,Y) = (0,1))| = 0,$$

хотя величины зависимы

4. а) Кривая $c(t)$ идет вдоль оси OX, значит нормаль к ней — вдоль OY. Отсюда $S = X$, $T = Y$, S распределено $R[-1, 1]$, $\mathbf{P}(T \in \cdot | S = s)$ будет также равномерным $R[-1, 1]$ распределением независимо от s.

б) В первом случае опять же $S = X$, $T = Y$. При этом $\mathbf{P}(S \leq x) = x^2/2$, $x \in [0, 1/2]$, $\mathbf{P}(S \leq x) = 1 - (1-x)^2/2$, $x \in [1/2, 1]$. Эти вычисления вытекают из простого подсчета площади множества $X \leq x$, поскольку X, Y распределены равномерно. $\mathbf{P}(T \leq x | S = s) = (x + 1 - s)/(2(1-s))$ при $x \in [s-1, 1-s]$, поскольку при фиксированном $S = s$ величина T распределена по $[s-1, 1-s]$ равномерно.

Во втором случае мы попадаем задачу 4а).

в) Поскольку параметр натуральный, $Ra = 1$. Исходя из соображений математического ожидания $E(T|S) = 0$. Но $T + R$ при условии S имеет то же распределение, что и расстояние $\tilde{R}$ от точек кольца до центра. Это распределение легко находится, поскольку $\mathbf{P}(\tilde{R} \leq x) = (x^2 - 1/4)/(3/4)$ при $x \in [1/2, 1]$. Значит, $\mathbf{E}(T|S) = \frac{8}{3}\int_{1/2}^1 x^2 dx - R = 7/9 - R$, т.е. $R = 7/9$, $a = 9/7$. S при этом распределен равномерно по $0, 14\pi/9$, $T|S = s$ имеет распределение с плотностью $8(x + 7/9)/3$ при $x \in [-5/18, 2/9]$.

5. Имеем

$$covLoc_{X,Y}(s) = -(\mathbf{D}S - \mathbf{D}T)\sin s \cos s,\ covGC(X,Y) = (\mathbf{D}S - \mathbf{D}T)/2\sqrt{E\sin^2(2S)},$$

где

$$\mathbf{E}\sin^2(2S) = \frac{1}{2\pi}\int_0^{2\pi} \sin^2(2s) ds = \frac{1}{2},\ \mathbf{D}S = \mathbf{E}S^2 - (\mathbf{E}S)^2 = \frac{\pi^2}{3},\ \mathbf{D}T = \frac{2}{(3/2)^2}\int_0^{3/2} t(t-1)^2 dt = \frac{1}{8}.$$

Как мы видим, ковариация ненулевая.