

Корреляция: простая, частная и условная

Борис Демешев*

27 апреля 2016 г.

Аннотация

Корреляция — это способ описать силу линейной зависимости между двумя случайными величинами одним числом. Каков геометрический смысл корреляции? Что такое частная корреляция? Как связаны частная и условная корреляция?

Ключевые слова: корреляция, частная корреляция, условная корреляция, косинус, проекция.

1 Сколько вешать в граммах?

Почему при болезни мы измеряем температуру тела? По всей видимости важны два факта:

1. Измерить температуру очень удобно
2. Это измерение несёт в себе информацию о здоровье

Сложное описание состояния здоровья сводится измерением температуры к одной цифре. Естественно, куча информации теряется в этой цифре и бессмысленно лечить человека, руководствуясь только температурой его тела. Температура 40° говорит, что что-то не так, но что — непонятно. А температура 36.6° ещё не говорит о том, что у человека идеальное здоровье. Однако процедура очень проста и в некоторых ситуациях, например, при обыкновенной простуде её достаточно для принятия решения о приёме жаропонижающего.

Если бы для измерения температуры нужно было специальное устройство размером с половину комнаты, никто бы дома её не мерял. Простота измерения очень важна!

Подобном образом дела обстоят и с описанием зависимости между случайными величинами. Зависимость между случайными величинами полностью описывается их совместной функцией распределения, $F(x, y) = \mathbb{P}(X \leq x, Y \leq y)$. Это слишком сложно! Вместо сложной функции распределения мы хотим получить одно число. Некую «силу зависимости». Назовём это мифическое число $Dep(X, Y)$.

Чего бы нам хотелось от этого числа?

1. Посчитать это число очень удобно
2. Это число несёт в себе информацию о зависимости

В случае двух случайных величин наиболее популярными мерами зависимости являются ковариация, корреляция и коэффициент в парной регрессии.

*НИУ ВШЭ, Москва.

2 Ковариация

В каком смысле «удобно» считать?

Очень часто возникают суммы случайных величин, поэтому было бы здорово, чтобы у суммы легко считалась наша характеристика $Dep(X + Z, Y)$. Проще всего было бы, если бы:

$$Dep(X + Z, Y) = Dep(X, Y) + Dep(Z, Y)$$

И, конечно, мы ждём, что у независимых случайных величин нулевая сила зависимости, $Dep(X, Y) = 0$, а ненулевая сила зависимости, $Dep(X, Y) \neq 0$, была бы возможна только у зависимых случайных величин.

Этим двум требуемым свойствам, простота подсчёта и информация о зависимости, отвечает ковариация.

Определение 1. Ковариация величин X и Y измеряет ...

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

здесь рассказать про прямоугольники и площадь с плюсом/минусом?

3 Коэффициент в парной регрессии

...

Недостаток коэффициента парной регрессии в его несимметричности. В регрессиях y на x и x на y коэффициенты не только не обязаны совпадать, но даже в произведении чаще всего не дают единицу.

От несимметричности коэффициент в парной регрессии можно избавить отмасштабировав y и x на их стандартные отклонения. При этом мы получим стандартизованный коэффициент парной регрессии.

...

Естественно, стандартизованный коэффициент связан с обычным соотношением:

$$\beta^{st} = \beta \cdot \frac{\sigma_X}{\sigma_Y}$$

По смыслу стандартизованный коэффициент показывает на сколько своих стандартных отклонений в среднем растёт величина Y при росте величины X на одно своё стандартное отклонение. И именно этот стандартизованный коэффициент парной регрессии и называется корреляцией.

4 Корреляция по-русски

Обычно в учебниках даётся такое определение корреляции

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}. \quad (4.1)$$

Естественно, возникает вопрос: «С какого перепугу? Почему это мы делим ковариацию на что-то там?»

Мы дадим ещё раз дадим определение корреляции словами:

Определение 2. Корреляция между случайными величинами X и Y показывает на сколько своих стандартных отклонений в среднем растёт случайная величина Y при росте случайной величины X на одно своё стандартное отклонение.

А теперь из этого словесного определения мы получим формулу ???. Разложим величину Y на два слагаемых. Первое слагаемое вбирает в себя всю ту часть Y , которая линейно зависит от X , а второе — всё оставшееся:

$$\frac{Y}{\sigma_Y} = \rho \cdot \frac{X}{\sigma_X} + \varepsilon$$

В этой формуле видно, что с ростом X на одно стандартное отклонений σ_X правая часть изменится в среднем на ρ , и, следовательно, величина Y в среднем изменится на $\rho \cdot \sigma_Y$.

Естественно, мы хотим, чтобы с ростом X величина ε в среднем не менялась, то есть хотим нулевую «силу зависимости» между ними, $\text{Cov}(X, \varepsilon) = 0$.

$$\text{Cov}\left(X, \frac{Y}{\sigma_Y} - \rho \cdot \frac{X}{\sigma_X}\right) = 0$$

По свойствам ковариации получаем

$$\text{Cov}(X, Y)/\sigma_Y - \rho \text{Cov}(X, X)/\sigma_X = 0$$

И, тадам, выражаем корреляцию, ρ :

$$\rho = \frac{\text{Cov}(X, Y)/\sigma_Y}{\text{Cov}(X, X)/\sigma_X} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Отметим асимметричность исходного разложения: эpsilon прибавляется в правой части уравнения к величине X . Несмотря на эту асимметричность результирующая формула для корреляции получается симметричной. Из этого следует, что ровно такой же результат получится, если начать с разложения:

$$\frac{X}{\sigma_X} = \rho \cdot \frac{Y}{\sigma_Y} + \varepsilon$$

Из определения неочевидно, что корреляция лежит в пределах от -1 до 1

Стоит обратить внимание на немного контр-интуитивный факт. Если бы зависимость между X и Y была бы жесткой детерминистической, и с ростом X на единицу величина Y росла бы на Δ , то с ростом Y на единицу величина X росла бы на $1/\Delta$. Для случайных величин обращения не происходит. Если с ростом X на одно своё стандартное отклонение величина Y в среднем растёт на ρ своих стандартных отклонений, то и с ростом Y на одно своё стандартное отклонение величина X в среднем растёт на ρ своих стандартных отклонений.

5 Геометрический смысл корреляции

Давайте рисовать случайные величины векторами-стрелочками! Не в том смысле, что у стрелочки случайное направление или длина, а в том смысле, что направление и длина стрелочки описывают характеристики этой случайной величины.

Любую геометрию можно задать, задав скалярное произведение. Действительно, если мы умеем считать скалярное произведение двух любых векторов, $\langle \vec{a}, \vec{b} \rangle$, то длина вектора считается ровно как в 9-м классе:

$$|\vec{a}| = \sqrt{\langle \vec{a}, \vec{a} \rangle}$$

И также любой девятиклассник помнит, что косинус угла между векторами считается как

$$\cos(\vec{a}, \vec{b}) = \frac{\langle \vec{a}, \vec{b} \rangle}{|\vec{a}| |\vec{b}|}$$

Мы определим скалярное произведение двух случайных величин как их ковариацию:

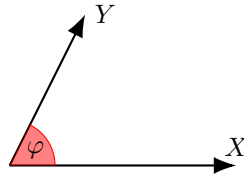
$$\langle X, Y \rangle = \text{Cov}(X, Y)$$

При таком подходе длиной случайной величины окажется стандартное отклонение:

$$\sqrt{\text{Cov}(X, X)} = \sqrt{\text{Var}(X)} = \sigma_X$$

А корреляция окажется косинусом угла между случайными величинами:

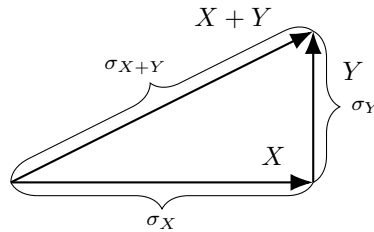
$$\cos \varphi = \cos(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \text{Corr}(X, Y)$$



Значит в нашей геометрии длина стрелочки — стандартное отклонение случайной величины, а косинус угла между двумя стрелочками — это корреляция двух случайных величин. Дисперсия, следовательно, это квадрат длины случайной величины. Перпендикулярными случайными величинами будут те, косинус угла между которыми равен нулю, то есть некоррелированные.

Например, сформулируем в данной геометрии теорему Пифагора. Если случайные величины X и Y перпендикулярны (корреляция или ковариация равны нулю), то дисперсия их суммы (квадрат длины гипотенузы) равен сумме их дисперсий (сумму квадратов длин катетов):

$$\text{Var}(X + Y) = \sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 = \text{Var}(X) + \text{Var}(Y)$$



Введение геометрии позволяет говорить о проекции. Например, можно спроецировать случайную величину Y на множество случайных величин пропорциональных величине X , $\{cX | c \in \mathbb{R}\}$. Если на обычной плоскости спроецировать вектор \vec{a} на прямую, порожденную вектором \vec{b} , то получится $\cos(\vec{a}, \vec{b}) \cdot \vec{b}$. По аналогии, если спроецировать случайную величину Y на множество $\{cX | c \in \mathbb{R}\}$, то получится $\text{Corr}(X, Y) \cdot X$. Другими словами, среди случайных величин пропорциональных X величина $\hat{Y} = \text{Corr}(X, Y) \cdot X$ — самая похожая на величину Y .

Понятие проекции позволяет интерпретировать квадрат корреляции. Квадрат косинуса равен отношению квадрата длины прилежащего катета $\text{Var}(\hat{Y})$ к квадрату гипотенузы $\text{Var}(Y)$.

(картинка)

Следовательно, $\text{Corr}(X, Y)^2 = \frac{\text{Var}(\hat{Y})}{\text{Var}(Y)}$, то есть квадрат корреляции показывает долю дисперсии Y , которую можно объяснить с помощью величин пропорциональных X .

6 Корреляция и независимость

Теорема 1. *Случайные величины X и Y независимы тогда и только тогда, когда некоррелированы любые функции $f(X)$ и $g(Y)$.*

Другими словами для независимости X и Y необходима некоррелированность пар X и Y , X^2 и $\cos(Y)$, $\exp(X)$ и $1/Y$, и так далее. Из этого следует, что некоррелированность X и Y является необходимым, но недостаточным условием для независимости.

Можно выделить три «степени» независимости случайных величин X и Y :

Некоррелированность Y и X	$\text{Corr}(X, Y) = 0$
$\mathbb{E}(Y X) = \mathbb{E}(Y)$	$\text{Corr}(f(X), Y) = 0$ для всех $f()$
Независимость Y и X	$\text{Corr}(f(X), g(Y)) = 0$ для всех $f()$ и $g()$

Многие ошибочно считают, что если величина X имеет нормальное распределение $N(\mu_X, \sigma_X^2)$ и величина Y имеет нормальное распределение $N(\mu_Y, \sigma_Y^2)$, и X и Y некоррелированы, то они независимы. Это неверно.

Контрпример. Случайная величина X имеет стандартное нормальное распределение $N(0; 1)$, случайная величина Z независима от X и равновероятно принимает значения -1 и $+1$. Определим величину Y как их произведение, $Y = XZ$.

В этом примере величины X и Y зависимы, так как $|X| = |Y|$. Однако Y распределена нормально стандартно и $\text{Corr}(X, Y) = 0$.

Правильная теорема звучит так:

Теорема 2. *Если некоррелированные случайные величины X и Y имеют совместное нормальное распределение, то X и Y независимы.*

Попутно упомянем ещё одно неожиданное свойство предъявленного контрпримера. Если случайные величины нормальны по отдельности, то вполне возможно, что их сумма ненормальна. Для пары величин, имеющих совместное нормальное распределение, это невозможно.

7 Более двух случайных величин

Если случайных оказывается больше двух, то для описания зависимости с помощью одного числа оказывается очень много вариантов. Среди популярных можно назвать:

1. коэффициент бета в множественной регрессии
2. стандартизованный коэффициент множественной регрессии
3. частная корреляция
4. условная корреляция
5. каноническая корреляция

8 Стандартизованный коэффициент регрессии

Для нахождения стандартизованного коэффициента множественной регрессии используется разложение

$$\frac{Y}{\sigma_Y} = \rho_{XY|Z} \cdot \frac{X}{\sigma_X} + \rho_{YZ|X} \frac{Z}{\sigma_Z} + \varepsilon$$

Два подхода к определению стандартизованного коэффициента регрессии эквивалентны в силу теоремы Фриша-Ву-Ловелла (Frisch–Waugh–Lovell). Обычно эта теорема формулируется применительно к регрессии, а здесь мы приведём её вариант для случайных величин.

Теорема 3. Если имеют место разложения:

$$Y = a_1 Z_1 + a_2 Z_2 + \dots + a_n Z_n + \tilde{Y}, \text{ где } \tilde{Y} \perp Z_1, Z_2, \dots, Z_n$$

и

$$X = b_1 Z_1 + b_2 Z_2 + \dots + b_n Z_n + \tilde{X}, \text{ где } \tilde{X} \perp Z_1, Z_2, \dots, Z_n$$

То в разложениях

$$\tilde{Y} = d\tilde{X} + \varepsilon, \text{ где } \varepsilon \perp \tilde{X}$$

и

$$Y = c_1 Z_1 + c_2 Z_2 + \dots + c_n Z_n + dX + u, \text{ где } u \perp Z_1, Z_2, \dots, Z_n, X$$

коэффициенты при \tilde{X} и X совпадают.

9 Частная корреляция

Определение 3. Частная корреляция между величинами X и Y при фиксированной величине Z показывает на сколько своих стандартных отклонений в среднем вырастет величина Y , очищенная от связи с Z , при росте величины X , очищенной от связи с Z , на одно своё стандартное отклонение.

Подход к подсчёту частной корреляции следующий:

1. Спроецируем X на множество величин, некоррелированных с Z . Получим \tilde{X} .
2. Спроецируем Y на множество величин, некоррелированных с Z . Получим \tilde{Y} .
3. Частная корреляция между X и Y при фиксированной Z — это обычная корреляция между \tilde{X} и \tilde{Y} .

(картинка ...)

10 Частная дисперсия и ковариация

Определения частной корреляции можно сформулировать с помощью частной ковариации и частной дисперсии.

$$pCorr(X, Y|Z) = \frac{pCov(X, Y|Z)}{\sqrt{pVar(X|Z)pVar(Y|Z)}}$$

Частная дисперсия величины X очищенной от Z определяется как

$$pVar(X|Z) = \text{Var}(X - aZ), \text{ где константа } a \text{ такова, что } \text{Cov}(X - aZ, Z) = 0$$

Частная ковариация

$$pCov(X, Y|Z) = \text{Cov}(X - aZ, Y - bZ), \text{ где константы } a \text{ и } b \text{ такова, что } \text{Cov}(X - aZ, Z) = \text{Cov}(Y - bZ, Z) = 0,$$

11 Условная корреляция

Определение 4. Условная корреляция между величинами X и Y при известном значении величины Z показывает на сколько своих стандартных отклонений σ_Y в среднем вырастет Y при росте величины X на одно своё стандартное отклонение σ_X при заданном значении величины Z .

Следует подчеркнуть одно существенное отличие условной корреляции от обычной и частной. Обычная и частная корреляция являются константами. Условная корреляция $\text{Corr}(X, Y|Z)$ является функцией от Z . Величина Z является случайной, поэтому и условная корреляция $\text{Corr}(X, Y|Z)$ является случайной величиной.

Здесь регрессионное определение???

Чуть более формальное определение:

Определение 5.

$$\text{Corr}(X, Y|Z) = \frac{\text{Cov}(X, Y|Z)}{\sqrt{\text{Var}(X|Z)\text{Var}(Y|Z)}},$$

где $\text{Cov}(X, Y|Z) = \mathbb{E}(XY|Z) - \mathbb{E}(X|Z)\mathbb{E}(Y|Z)$ и $\text{Var}(X|Z) = \mathbb{E}(X^2|Z) - (\mathbb{E}(X|Z))^2$

Пример подсчета частной и условной корреляций.

Пример 1.

Закон распределения случайных величин X_1, X_2, X_3 задан двумя таблицами:

$X_3 = 0$			$X_3 = 1$		
	$X_2 = 0$	$X_2 = 1$		$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	0.06	0.08	$X_1 = 0$	0.1	0
$X_1 = 1$	0.24	0.32	$X_1 = 1$	0	0.2

Найдите условную корреляцию $\text{Corr}(X_1, X_2|X_3)$ и частную корреляцию $p\text{Corr}_{X_3}(X_1, X_2)$.

Решение.

Эти две таблички на самом деле реализуют простую мысль: при $X_3 = 1$ величины X_1 и X_2 связаны детерминистически линейно, а при $X_3 = 0$ величины X_1 и X_2 независимы.

Считаем две вспомогательные условные корреляции, $\text{Corr}(X_1, X_2|X_3 = 0) = 0$, $\text{Corr}(X_1, X_2|X_3 = 1) = 1$.

Отсюда получаем, что $\text{Corr}(X_1, X_2|X_3) = X_3$. Для дискретных случайных величин запись условного ожидания не однозначна, и, например, ответ $\text{Corr}(X_1, X_2|X_3) = X_3^2$ также будет верным.

Пример 2.

Величины X_1, X_2, X_3 имеют совместное нормальное распределение с математическим ожиданием $\mathbb{E}(X) = (1, 2, -3)'$ и ковариационной матрицей

$$\text{Var}(X) = \begin{pmatrix} 9 & 2 & -1 \\ 2 & 16 & 1 \\ -1 & 1 & 4 \end{pmatrix}$$

Найдите условную корреляцию $\text{Corr}(X_1, X_2|X_3)$ и частную корреляцию $\text{rCorr}_{X_3}(X_1, X_2)$.

Решение.

...

В данном примере частная и условная корреляция совпали. Это одно из приятных свойств многомерного нормального распределения:

Теорема 4. Если величины X, Y и Z имеют совместное нормальное распределение, то:

1. частная и условная дисперсии совпадают
2. частная и условная ковариации совпадают
3. частная и условная корреляции совпадают

12 Случай равных дисперсий

Если дисперсии всех рассматриваемых (или достаточно только коррелируемых?) величин равны, то частная корреляция Y и X при фиксированном Z равна стандартизованному коэффициенту регрессии, равна обычному коэффициенту регрессии Y на X и остальные, равна обычному коэффициенту регрессии X на Y и остальные.

К данному случаю относиться частная автокорреляционная функция для стационарных процессов:

Пример 3. AR(1) процесс

13 Каноническая корреляция

Если разбить все случайные величины на две группы, скажем на иксы и игреки, то каноническая корреляция — это максимальная корреляция между линейной комбинацией иксов и линейной комбинацией игреков.

$$\text{canCorr}(Y_1, Y_2, Y_3; X_1, X_2) = \max_{\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3} \text{Corr}(\alpha_1 X_1 + \alpha_2 X_2, \beta_1 Y_1 + \beta_2 Y_2 + \beta_3 Y_3)$$

Если группа игроков состоит из одной случайной величины, то мы получаем коэффициент множественной корреляции, квадрат которого можно назвать «теоретическим R^2 ».

$$canCorr(Y; X_1, X_2, \dots, X_k)$$

И Шахерезаду застигло утро, и она прекратила дозволенные речи, заметив, что вышла замечательная статья Димы Борзых о методе канонических корреляций.

14 Матрица — Мать всех регрессий и корреляций

Оказывается многие упомянутые меры зависимости порождаются обратной ковариационной матрицей.

Пусть X — это вектор случайных величин $X = (X_1, X_2, \dots, X_k)'$ с ковариационной матрицей $C = \text{Var}(X)$. Тогда обратная ковариационная матрица имеет следующую структуру:

$$C^{-1} = \begin{pmatrix} \frac{1}{p\text{Var}(X_1|X_{-1})} & & \\ & \frac{1}{p\text{Var}(X_2|X_{-2})} & \\ & & \frac{1}{p\text{Var}(X_3|X_{-3})} \end{pmatrix}$$

Похожий результат верен и для обратной корреляционной матрицы. Если A — корреляционная матрица, то:

$$A^{-1} = \begin{pmatrix} \frac{\text{Var}(X_1)}{p\text{Var}(X_1|X_{-1})} & & \\ & \frac{\text{Var}(X_2)}{p\text{Var}(X_2|X_{-2})} & \\ & & \frac{\text{Var}(X_3)}{p\text{Var}(X_3|X_{-3})} \end{pmatrix}$$

Частная корреляция между X_i и X_j при фиксированных остальных иксах равна:

$$pCorr(X_i, X_j | \dots) = -\frac{A_{ij}}{\sqrt{A_{ii}A_{jj}}}$$

Стандартизованный коэффициент при X_j в регрессии X_i на остальные иксы равен:

$$\hat{\beta}_{ij}^{st} = -\frac{A_{ij}}{A_{ii}}$$

Квадрат коэффициента множественной корреляции X_i с остальными иксами равен:

$$canCorr(X_i; \dots) = 1 - \frac{1}{A_{ii}}$$

Доказательство:

15 Частные корреляции бывают любые

Произвольная матрица не может быть корреляционной матрицей вектора случайных величин. Она должна быть симметричной и неотрицательно определённой, а на диагонали должны стоять единицы.

Однако произвольная симметричная матрица из чисел в диапазоне $(-1; 1)$ может быть частной корреляционной матрицей. Более того, у стационарного процессам

может быть совершенная любая частная автокорреляционная функция со значениями в диапазоне $(-1; 1)$. Например, у стационарного процесса может быть частная автокорреляционная функция

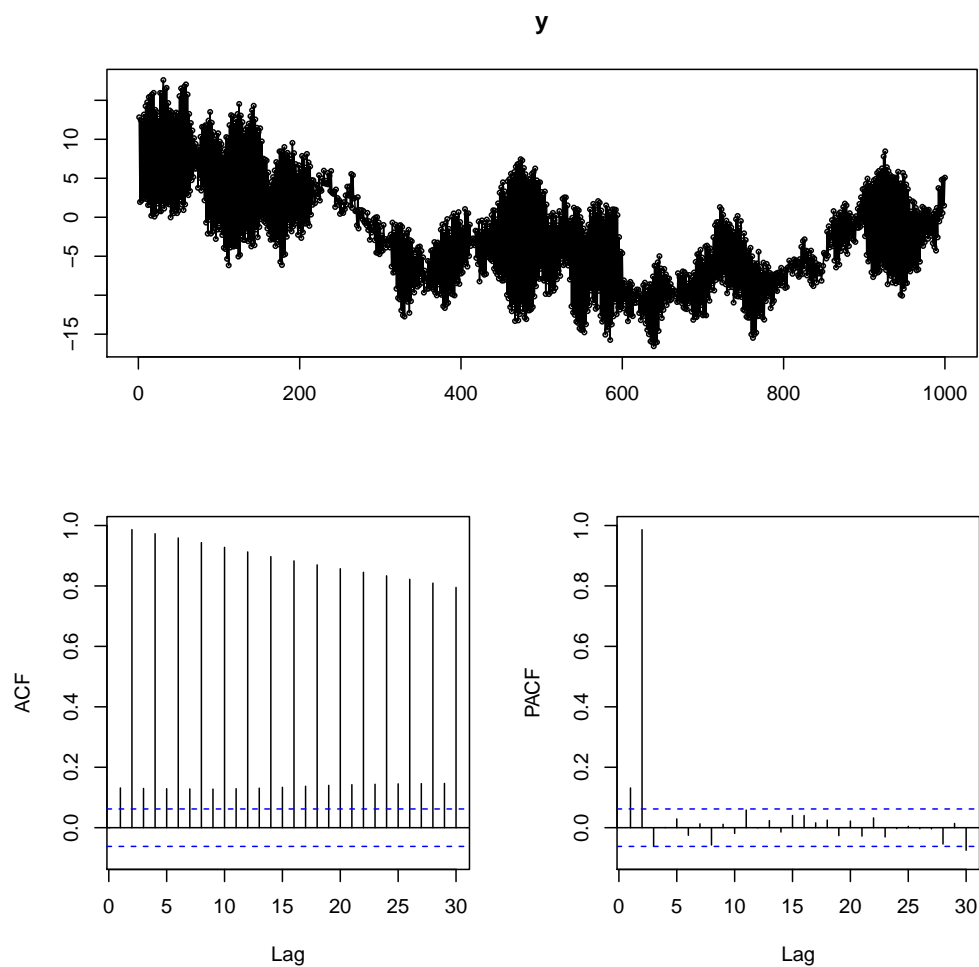
$$\varphi_{kk} = \begin{cases} 0, & \text{если } k = 1 \\ 0.99, & \text{если } k = 2 \\ 0, & \text{если } k \geq 3 \end{cases}$$

А вот и он, $y_t = 0.99y_{t-2} + \varepsilon_t$:

```
library("forecast")

## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric
## Loading required package: timeDate
## Loading required package: methods
## This is forecast 6.2

y <- arima.sim(n = 1000, model = list(ar = c(0, 0.99)))
tsdisplay(y)
```



16 Почташки

https://en.wikipedia.org/wiki/Schur_complementApplications_to_probability_theory_and_statistics

<https://stats.stackexchange.com/questions/140080/why-does-inversion-of-a-covariance-matrix-yield-partial-correlations-between-ran>

<https://stats.stackexchange.com/questions/10795/how-to-interpret-an-inverse-covariance-or-precision-matrix>

<http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1147context=ejsie>

<https://cran.r-project.org/web/packages/corpcor/corpcor.pdf>

17 Выборочные характеристики

это в статью два?

В теории обычную корреляцию и частную корреляцию можно посчитать, если известен закон распределения случайных величин. На практике закон распределения не известен, однако доступны наблюдения. Как по имеющимся наблюдениям оценить неизвестные корреляции?

Несколько способов оценки корреляции

Здесь пара картинок: википедийная с корреляциями и два ряда случайного блуждания/тренда

Несколько способов оценки частной корреляции

Оценкой для частной дисперсии X_1 при фиксированных X_2, \dots, X_n является оценка дисперсии случайной составляющей в регрессии X_1 на X_2, \dots, X_n

Здесь про зависимости возникающие при оценке автокорреляционной и частной автокорреляционной функции для случайных процессов.

Несколько способов оценки множественной корреляции:

Если $\text{canCorr}(Y; X_1, X_2, \dots, X_k) = 0$, то $\mathbb{E}(R_{adj}^2) = 0$.

Обратная к выборочной ковариационной матрице

$$\hat{C}^{-1} = \begin{pmatrix} \frac{1}{\hat{\sigma}_1^2} & \frac{1}{\hat{\sigma}_2^2} & \frac{1}{\hat{\sigma}_3^2} \end{pmatrix}$$

Если мы предположим, что разные наблюдения в выборке независимы и одинаково распределены, то получим такие выборочные оценки

Теоретический	Выборочный
$pVar(y_i z_i)$	$\hat{\sigma}_u^2$ в регрессии y на z , выборочная дисперсия остатков
$pCov(y_i, x_i z_i)$	выборочная ковариация остатков двух регрессий
$\text{canCorr}(y_i; x_{i1}, x_{i2}, \dots, x_{ik})$	R^2, R_{adj}^2
коэффициент в множественной регрессии	$\hat{\beta} = (X'X)^{-1}X'y$
$\hat{\beta}_j^{st} = \hat{\beta}_j \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2}, \hat{\beta}^{st} = (X^{st'}X^{st})^{-1}X^{st'}y^{st}$	несмещенный?
обычная корреляция	$sCorr(y, x) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\dots}}$
частная корреляция	корреляция остатков регрессии, обращение корреляционной