



ФИНАНСОВЫЙ
УНИВЕРСИТЕТ

ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПАРАМЕТРИЗАЦИЯ ЗАДАЧ ПО ДИСЦИПЛИНЕ «ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА» НА ПЛАТФОРМЕ MOODLE С ИСПОЛЬЗОВАНИЕМ ИНСТРУМЕНТАРИЯ JUPYTER NOTEBOOK

Золотаревский Артём, 2 курс, бакалавр
artyom@zolotarevskiy.ru

Научный руководитель:
П.Е. Рябов, д.ф. - м.н., проф.
PERyabov@fa.ru

1. Формулировка проблемы и цели работы
2. Ключевые результаты
3. Описание рабочего процесса
4. Примеры параметризованных задач
5. Преимущества предложенного решения
6. Текущие задачи
7. Список литературы

Проблема: качественная подготовка практических и оценочных материалов является отнюдь непростой задачей, занимающей существенную часть преподавательского времени. При этом важно обеспечить самостоятельность выполнения работ студентами, не допустить списывания вариантов других учащихся.

Цель работы: разработать программное решение, автоматизирующее процесс подготовки практических задач, для решения описанной проблемы.

Разработан программный модуль «*TaskGenerator*» для параметризации задач математических дисциплин, который:

- базируется на отлично зарекомендовавших себя и проверенных временем инструментах;
- объединяет несколько технологий в единый программный конвейер, сохраняя при этом простоту использования;
- предоставляет удобный интерфейс, не ограничивая пользователя в функциональных возможностях.

Что позволяет делать новый программный продукт?

- Позволяет генерировать в автоматическом режиме неограниченное количество уникальных вариантов задач.
- Позволяет объединять сгенерированные задачи в виде уникальных вариантов раздаточных материалов, самостоятельных, контрольных, экзаменационных билетов, иным языком говоря, по заданному шаблону для любых целей образовательной деятельности.
- Доступна выгрузка созданных вариантов в систему управления обучением «Moodle».

Что программа не делает?

Разработка задач и наполнение соответствующей базы данных этими задачами по-прежнему лежит на плечах преподавателя. При этом требуется достаточный навык владения ключевыми технологиями.

TEX



python™
3



Jupyter

Система управления образовательными электронными курсами «Moodle».

В 2023 году база зарегистрированных пользователей включает более 350 миллионов по всему миру.



Вопрос 1

Пока нет
ответа

Балл: 12,00

1. (10) Признак $X(k)$ задан на множестве $\Omega = \{1, 2, \dots, 13\}$ следующей таблицей:

k	1	2	3	4	5	6	7	8	9	10	11	12	13
$X(k)$	2	5	3	5	5	4	5	6	5	4	3	5	5

Из Ω извлекается случайная бесповторная выборка объема 6. Найдите математическое ожидание и дисперсию среднего значения \bar{X} признака X в выборке.

Ответ введите с точностью до трех знаков после запятой.

Математическое ожидание $\mathbb{E}(\bar{X}) =$

Дисперсия $\text{Var}(\bar{X}) =$

1. (10) Признак $X(k)$ задан на множестве $\Omega = \{1, 2, \dots, 13\}$ следующей таблицей:

k	1	2	3	4	5	6	7	8	9	10	11	12	13
$X(k)$	2	5	3	5	5	4	5	6	5	4	3	5	5

Из Ω извлекается случайная бесповторная выборка объема 6. Найдите математическое ожидание и дисперсию среднего значения \bar{X} признака X в выборке.

Ответ введите с точностью до трех знаков после запятой.

Математическое ожидание $\mathbb{E}(\bar{X}) =$ ✓

Один из возможных правильных ответов: 4,38462

Дисперсия $\text{Var}(\bar{X}) =$ ✓

Один из возможных правильных ответов: 0,11275

Задача 1 ("Метод максимального правдоподобия"): По содержащейся в файле `dataFileName` реализации случайной выборки из двумерного нормального распределения

$N\left(\mathbb{E}(X); \mathbb{E}(Y); \begin{vmatrix} \text{Var}(X) & \text{Cov}(X,Y) \\ \text{Cov}(X,Y) & \text{Var}(Y) \end{vmatrix}\right) = N\left(\text{mu}X; \text{mu}Y; \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ с неизвестными параметрами $\rho \in (-1; 1)$ и $\sigma > 0$: 1) запишите логарифм функции правдоподобия, $\ln L(\rho, \sigma)$; 2) найдите оценки максимального правдоподобия $\hat{\rho}$ и $\hat{\sigma}$.

Решение:

1) Уравнение максимального правдоподобия для модели $N\left((a; b); \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ имеет вид

$$\ln L = -\frac{A_0 + B_0 - 2C_0\rho}{2(1 - \rho^2)\sigma^2} - \ln(2\pi) - \frac{1}{2} \ln(1 - \rho^2) - \ln(\sigma^2)$$

где $A_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2$; $B_0 = \frac{1}{n} \sum_{k=1}^n (y_k - b)^2$; $C_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)(y_k - b)$;

2) Система для определения оценок имеет вид:

$$\begin{cases} \frac{\partial \ln L}{\partial \rho} = \frac{\rho}{1-\rho^2} - \frac{\rho(A_0+B_0-2C_0\rho)}{(1-\rho^2)^2\sigma^2} + \frac{C_0}{(1-\rho^2)\sigma^2} = 0 \\ \frac{\partial \ln L}{\partial \sigma} = \frac{A_0+B_0-2C_0\rho}{(1-\rho^2)\sigma^3} - \frac{2}{\sigma} = 0, \end{cases}$$

Откуда

$$\hat{\rho} = \frac{2C_0}{A_0+B_0} = \text{hatRho}$$
$$\hat{\sigma} = \sqrt{\frac{A_0+B_0}{2}} = \text{hatSigma}$$

Ответ:

- $\hat{\rho} = \frac{2C_0}{A_0+B_0} = \text{hatRho}$;
- $\hat{\sigma} = \sqrt{\frac{A_0+B_0}{2}} = \text{hatSigma}$.

О рабочем процессе

Jupyter MaximumLikelihoodMethod Последняя контрольная точка: 16 минут назад (несохраненные изменения)



Разлогиниться

Файл Редактировать Вид Вставка Ячейка Ядро Widgets LaTeX_envs Справка

Не доверять

Python 3 (ipykernel) O

Код

Ввод [42]: 1 taskgen.show(variables)

```
10:14:49: Подставляем переданные значения в шаблон "/Users/artiomzlotarevskiy/Documents/university/ryabov/taskgen/bank/Q2/theme_1/MaximumLikelihoodMethod/MaximumLikelihoodMethod.tex"...
10:14:49: Конвертируем файл /Users/artiomzlotarevskiy/Documents/university/ryabov/taskgen/bank/Q2/theme_1/MaximumLikelihoodMethod/tmp/subs_16.03.2023_10-14-49.tex...
10:14:52: Удаляем временные файлы для /Users/artiomzlotarevskiy/Documents/university/ryabov/taskgen/bank/Q2/theme_1/MaximumLikelihoodMethod/tmp/subs_16.03.2023_10-14-49.tex...
```

```
[STATUS] make4ht: Conversion started
[STATUS] make4ht: Input file: subs_16.03.2023_10-14-49.tex
[STATUS] make4ht: Conversion finished
```

Out[42]:

Задача 1 ("Метод максимального правдоподобия"): По содержащейся в файле ds-mlm-0.csv реализации случайной выборки из двумерного нормального распределения

$N\left(\left(\mathbb{E}(X); \mathbb{E}(Y)\right); \begin{vmatrix} \text{Var}(X) & \text{Cov}(X,Y) \\ \text{Cov}(X,Y) & \text{Var}(Y) \end{vmatrix}\right) = N\left(-2; 4; \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ с неизвестными параметрами $\rho \in (-1; 1)$ и $\sigma > 0$: 1) запишите логарифм функции правдоподобия, $\ln L(\rho, \sigma)$; 2) найдите оценки максимального правдоподобия $\hat{\rho}$ и $\hat{\sigma}$.

Решение:

1) Уравнение максимального правдоподобия для модели $N\left((a; b); \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ имеет вид

$$\ln L = -\frac{A_0 + B_0 - 2C_0\rho}{2(1 - \rho^2)\sigma^2} - \ln(2\pi) - \frac{1}{2} \ln(1 - \rho^2) - \ln(\sigma^2)$$

где $A_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2$; $B_0 = \frac{1}{n} \sum_{k=1}^n (y_k - b)^2$; $C_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)(y_k - b)$;

2) Система для определения оценок имеет вид:

$$\left\{ \begin{aligned} \frac{\partial \ln L}{\partial \rho} &= \frac{\rho}{1 - \rho^2} - \frac{\rho(A_0 + B_0 - 2C_0\rho)}{(1 - \rho^2)^2 \sigma^2} + \frac{C_0}{(1 - \rho^2)\sigma^2} = 0 \end{aligned} \right.$$



добоодбия $\hat{\rho}$ и $\hat{\sigma}$.

Решение:

1) Уравнение максимального правдоподобия для модели $N\left((a; b); \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ имеет вид

$$\ln L = -\frac{A_0 + B_0 - 2C_0\rho}{2(1 - \rho^2)\sigma^2} - \ln(2\pi) - \frac{1}{2}\ln(1 - \rho^2) - \ln(\sigma^2)$$

где $A_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2$; $B_0 = \frac{1}{n} \sum_{k=1}^n (y_k - b)^2$; $C_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)(y_k - b)$;

2) Система для определения оценок имеет вид:

$$\begin{cases} \frac{\partial \ln L}{\partial \rho} = \frac{\rho}{1 - \rho^2} - \frac{\rho(A_0 + B_0 - 2C_0\rho)}{(1 - \rho^2)^2\sigma^2} + \frac{C_0}{(1 - \rho^2)\sigma^2} = 0 \\ \frac{\partial \ln L}{\partial \sigma} = \frac{A_0 + B_0 - 2C_0\rho}{(1 - \rho^2)\sigma^3} - \frac{2}{\sigma} = 0, \end{cases}$$

Откуда


$$\hat{\rho} = \frac{2C_0}{A_0 + B_0} = 0,506$$

$$\hat{\sigma} = \sqrt{\frac{A_0 + B_0}{2}} = 1,308$$

Ответ:

- $\hat{\rho} = \frac{2C_0}{A_0 + B_0} = 0,506$;
- $\hat{\sigma} = \sqrt{\frac{A_0 + B_0}{2}} = 1,308$.

🔒 pypi.org/project/taskgen/ 📄 ☆ ⚙️ 👤 ⚙️

 🔍

[Помощь](#) [Спонсоры](#) [Войти](#) [Зарегистрироваться](#)

taskgen 0.3.7.1

`pip install taskgen` 📄

✓ [Последняя версия](#)

Выпущен: 12 мар. 2023 г.

Ядро генератора банка задач на базе MiKTeX и Jupyter Notebook

Навигация

- ☰ [Описание проекта](#)
- 🕒 [История выпусков](#)
- 📄 [Загрузка файлов](#)

Ссылки проекта

- 🏠 [Homepage](#)

Описание проекта

Task Generator Core

Ядро генератора банка задач на базе MiKTeX и Jupyter Notebook.

Для работы с ядром необходимо использовать панель управления, которая размещена здесь - <https://github.com/artiom-zolotarevskiy/taskgen>

По всем вопросам: artiom@zolotarevskiy.ru ✉️

Возможности

О рабочем процессе

github.com/artiom-zolotarevskiy/taskgen

Search or jump to... Pull requests Issues Codespaces Marketplace Explore

artiom-zolotarevskiy / taskgen Public

Pin Unwatch 1 Fork 0 Star 0

<> Code Issues 3 Pull requests Actions Projects Wiki Security Insights Settings

main 2 branches 0 tags

Go to file Add file Code

Your main branch isn't protected
Protect this branch from force pushing or deletion, or require status checks before merging. [Learn more](#) [Protect this branch](#)

artiom-zolotarevskiy removed unnecessary logs 15c0be4 3 days ago 85 commits

bank	new core version	last week
settings	added ppitest block	3 days ago
templates/default	removed unnecessary logs	3 days ago
.gitignore	The core is separated from the control panel	3 weeks ago
README.md	Updated the command to install the kernel package	4 days ago
control_panel.ipynb	Added a command to get kernel updates	last week

☰ README.md ✎

Task Generator Control Panel

Панель управления генератором банка задач на базе MiKTeX и Jupyter Notebook.

About

Генератор экзаменационных билетов на базе MikTeX и Python

Readme

0 stars

1 watching

0 forks

Releases

No releases published

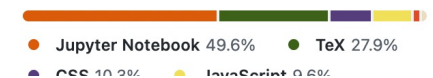
[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages



Билет № 3

1. (10) Признак $X(k)$ задан на множестве $\Omega = \{1, 2, \dots, 18\}$ следующей таблицей:

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$X(k)$	2	5	5	5	2	2	5	3	6	6	6	4	4	2	3	4	4	2

Из Ω извлекается случайная бесповторная выборка объема 5. Найдите математическое ожидание и дисперсию среднего значения \bar{X} признака X в выборке.

Ответ введите с точностью до трех знаков после запятой.

2. (10) Дайте определение случайной величины, которая имеет χ^2 -распределение с n степенями свободы. Запишите плотность χ^2 -распределения. Выведите формулы для математического ожидания $\mathbb{E}(X)$ и дисперсии $\text{Var}(X)$ хи-квадрат распределения с n степенями свободы. Найдите: а) $\mathbb{P}(\chi_{19}^2 > 20,812)$, где χ_{19}^2 – случайная величина, которая имеет χ^2 -распределение с 19 степенями свободы; б) 94% (верхнюю) точку $\chi_{0,94}^2(7)$ хи-квадрат распределения с 7 степенями свободы.

3. (10) Корзина содержит 24 шаров(а): 5 белых(й), 4 синих(й) и 15 красных(й). На каждом шаге из корзины «наудачу» извлекается шар и назад в корзину не возвращается. Исход n последовательных извлечений называется выборкой объема n **без возвращения** или **бесповторной выборкой**. Из корзины наудачу (без возвращения) извлечены 8 шаров. Найдите вероятности следующих событий и покажите их статистическую устойчивость:

- $A = \{\text{все шары красные}\};$
- $B = \{2 \text{ синих, } 1 \text{ белых и } 5 \text{ красных(й)}\};$
- $D = \{\text{в точности } 2 \text{ белых шара}\}.$
- $E = \{\text{в точности } 5 \text{ белых шаров(а) и по крайней мере } 6 \text{ красных шаров(а)}\}.$

4. (10) Случайная величина X (срок службы изделия) имеет распределение “экспоненциального типа” (the shifted exponential distribution), плотность которого задается формулой $f(x) = \lambda e^{-\lambda(x-\tau)}$, $x \geq \tau$, $\lambda > 0$. В таблице приведены сгруппированные данные по срокам службы (в часах) для $n = 205$ изделий:

Δ_i	[2,54; 20,54)	[20,54; 38,54)	[38,54; 56,54)	[56,54; 74,54)	[74,54; 92,54)	[92,54; 110,54)	[110,54; 146,54)
n_i	91	49	24	15	11	7	2

Найдите методом моментов точечные оценки неизвестных параметров λ и τ “экспоненциального” распределения. Используя полученные оценки, оцените время, которое изделие прослужит с вероятностью 82,93%.

Примеры параметризованных задач

1. (10) Случайные величины X и Y независимы и имеют равномерное распределение на отрезке $[-5; 14]$. Для случайной величины $Z = X + Y$ найдите:

- 1) функцию распределения $F_Z(x)$;
- 2) плотность распределения $f_Z(x)$ и постройте график плотности;
- 3) значение $F_Z(26,86)$.

Ответ: 1) Функция распределения $F_Z(x)$ имеет вид:
$$F_Z(x) = \begin{cases} 0, & x \leq -10; \\ \frac{(x+10)^2}{722}, & -10 \leq x \leq 9; \\ 1 - \frac{(x-28)^2}{722}, & 9 \leq x \leq 28; \\ 1, & x \geq 28; \end{cases}$$

2) Плотность распределения $f_Z(x)$ имеет вид:
$$f_Z(x) = \begin{cases} \frac{x}{361} + \frac{10}{361}, & -10 \leq x \leq 9; \\ \frac{28}{361} - \frac{x}{361}, & 9 \leq x \leq 28; \\ 0, & x \notin [-10; 28]; \end{cases}$$

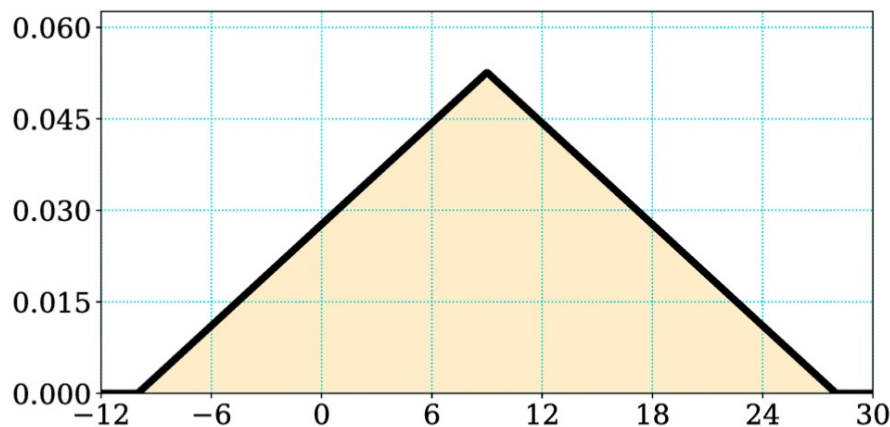


Рис. График плотности

3) Значение функции равно: $F_Z(26,86) = 0,998200$.

Примеры параметризованных задач

2. (10) По содержащейся в файле ds.1.2.1.csv реализации случайной выборки из двумерного нормального распределения

$N\left(\left(\mathbb{E}(X); \mathbb{E}(Y)\right); \begin{vmatrix} \text{Var}(X) & \text{Cov}(X,Y) \\ \text{Cov}(X,Y) & \text{Var}(Y) \end{vmatrix}\right) = N\left(-3; -1; \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ с неизвестными параметрами $\rho \in (-1; 1)$ и $\sigma > 0$:

1) запишите логарифм функции правдоподобия, $\ln L(\rho, \sigma)$; 2) найдите оценки максимального правдоподобия $\hat{\rho}$ и $\hat{\sigma}$.

Решение: 1) Уравнение максимального правдоподобия для модели $N\left((a; b); \begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}\right)$ имеет вид

$$\ln L = -\frac{A_0 + B_0 - 2C_0\rho}{2(1 - \rho^2)\sigma^2} - \ln(2\pi) - \frac{1}{2}\ln(1 - \rho^2) - \ln(\sigma^2)$$

где $A_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2$; $B_0 = \frac{1}{n} \sum_{k=1}^n (y_k - b)^2$; $C_0 = \frac{1}{n} \sum_{k=1}^n (x_k - a)(y_k - b)$;

2) Система для определения оценок имеет вид:

$$\begin{cases} \frac{\partial \ln L}{\partial \rho} = \frac{\rho}{1 - \rho^2} - \frac{\rho(A_0 + B_0 - 2C_0\rho)}{(1 - \rho^2)^2\sigma^2} + \frac{C_0}{(1 - \rho^2)\sigma^2} = 0 \\ \frac{\partial \ln L}{\partial \sigma} = \frac{A_0 + B_0 - 2C_0\rho}{(1 - \rho^2)\sigma^3} - \frac{2}{\sigma} = 0, \end{cases}$$

Откуда

$$\hat{\rho} = \frac{2C_0}{A_0 + B_0} = -0,385$$
$$\hat{\sigma} = \sqrt{\frac{A_0 + B_0}{2}} = 1,022$$

Ответ:

- $\hat{\rho} = \frac{2C_0}{A_0 + B_0} = -0,385$;
- $\hat{\sigma} = \sqrt{\frac{A_0 + B_0}{2}} = 1,022$.

- новая система позволяет генерировать задачи любой сложности дисциплины, включая различные распределения как дискретных, так и непрерывных случайных величин, а также выборки из многомерных распределений для их последующего статистического анализа;
- решает проблему корректного отображения математических шрифтов и переноса слов в разных типах файлов;
- упрощает процесс подготовки задач благодаря возможности использования «*Jupyter Notebook*» и профессионального редактора «*TeX*», такого как «*WinEdt*» или «*TeXstudio*»;
- предоставляет возможность определять собственные шаблоны создаваемых билетов;
- позволяет определять банк тестов с любыми названиями и уровнем иерархической вложенности папок и файлов;
- создав задачу единожды, она определена теперь раз и навсегда, происходит накопление данных, нет необходимости возвращаться и снова проводить работу, которая уже была до этого сделана.

- Наполнение банка задачами по дисциплине «Теория вероятности и математическая статистика»;
- Выявление новых сценариев, требующих расширения функционала системы;
- Разработка новых полезных функций как для «TeX», так и для «Python», упрощающих процесс подготовки задач;
- Публикация примеров сюжетов, использующих датасеты и изображения;
- Активное тестирование, устранение выявленных неточностей в работе;
- Расширение документации;
- Расширение функциональных возможностей работы с «Moodle».

- [1] Страница предложенного автором генератора банка задач TaskGenerator: [Электронный ресурс]. URL: <https://github.com/artiom-zolotarevskiy/taskgen>. (Дата обращения: 06.02.2023).
- [2] Кузнецов А. В. *Основы LaTeX* / А. В. Кузнецов // Национальный исследовательский ядерный университет «МИФИ», 2021. 364 с. Москва
- [3] Cervone D. P. *MathJax: A Platform for Mathematics on the Web* / D. P. Cervone // Notices of the American Mathematical Society, 2012.
- [4] Плагин вопросов qtype_formulas, поддерживающий случайные величины и несколько полей для ввода ответов: [Электронный ресурс]. URL: https://moodle.org/plugins/qtype_formulas. (Дата обращения: 06.02.2023).
- [5] Franklin M. L. *Word Hy-phen-a-tion by Com-put-er: dissertation* / M. L. Franklin // Department of Computer Science, Stanford University. – 1983. – Report No. STAN-CS-83-977.