

Четвертий етап полягає в дослідженні відхилень фактичних значень рівнів ряду від розрахункових рівнів тренду, а також у вимірі та моделюванні сезонних (циклічних) коливань.

П'ятим етапом є розрахунок прогнозованих значень часового ряду для майбутніх періодів та оцінка якості моделі.

Для оцінки якості прогнозу використовують середню абсолютну процентну помилку (Mean Absolute Percentage Error):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|E_i|}{y_i} * 100\% , \quad (3)$$

де E_i – різниця між фактичним і прогнозованим значенням.

Прогноз є задовільним якщо MAPE складає більше 50%.

Таким чином, було отримано прогнозу математичну модель для визначення показників на майбутній період за допомогою часових рядів. Перевага цього методу полягає в знаходженні прихованих короткострокових і довгострокових закономірностей, які впливають на досліджувану величину, що дозволяє застосовувати його в досить широких діапазонах. Точність такого прогнозу в багатьох випадках перевищує точність інших методів прогнозування.

Список літератури

1. Аскинадзи В.М. *Інвестиції: учебник для бакалавров* / В.М. Аскинадзи, В.Ф. Максимова - М.: Издательство Юрайт, 2014. – 112 с.
2. Архипова М.Ю. *Анализ данных: Учебно-метод. пособие* / М.Ю. Архипова – Мн.: БГЭУ, 2016. – 19с.
3. Афанасьев В.Н. *Анализ временных рядов и прогнозирование* / М.М. Юзбашев, В.Н. Афанасьев. М.: Финансы и статистика, 2017. — 228 с.

УДК 004.82:004:85

*Зінченко Б. В., студент 2 курсу
спеціальності 122 «Комп'ютерні науки»
Нескородєва Т. В., к.т.н., доцент, доцент
кафедри інформаційних технологій*

АНАЛІЗ ДАНИХ ПРО КОРОНОВІРУС У СВІТІ МЕТОДАМИ СТАТИСТИЧНОГО НАВЧАННЯ

Донецький національний університет імені Василя Стуса, м. Вінниця

Людство зустрілося з жахливою хворобою яка кожен день забирає безліч життів. Коронавіруси (лат. Coronaviridae) – сукупність одноланцюгових РНК-вірусів, що включає на січень 2020 року 39 видів вірусів, об'єднаних у дві підмножини Letovirinae та Orthocoronavirinae (до яких належить і SARS-CoV-2, який спричинив, згідно з міжнародною медичною термінологією, спалах

коронавірусної хвороби, що почався в грудні 2019). Віруси цієї множини уражають людину, котів, птахів, собак, велику рогату худобу, свиней, кажанів, деяких диких хижих ссавців тощо. Тому є актуальним дослідження набору даних про хворих на коронавірус у світі.

Його актуальність обумовлена наступними факторами: зростання хворих, також зростанням смертності. За даними на 20.04.2020 всього заражених 2 444 209 осіб, летальні випадки 168 тисяч, вилікувалося 640 252, зараз хворіє 1 635 971 де 56 238 у критичному стані.

Аналіз проведемо за допомогою пакету R. Для дослідження був використаний набір даних COVID-19 geographic-disbtributi, що містить 10537 спостережень по 9 показниках:

- day – день коли відбувся збір даних,
- month – місяць коли відбувся збір даних,
- year – рік коли відбувся збір даних,
- cases – кількість випадків зараження,
- deaths – кількість летальних випадків,
- countriesAndTerritories – країна,
- geoId – Id локації,
- countryterritoryCode – код країни,
- popData2018 – кількість населення на 2018 рік.

Побудувавши графік залежності (рис.1) смертності від випадків зареєстрованих хворих на коронавірус ми зробимо висновок, що поки що мало країн де більше 5000 тисяч хворих, і більше 5000 тисяч смертей але динаміки змін показують, що захворюваність і смертність буде збільшуватись.

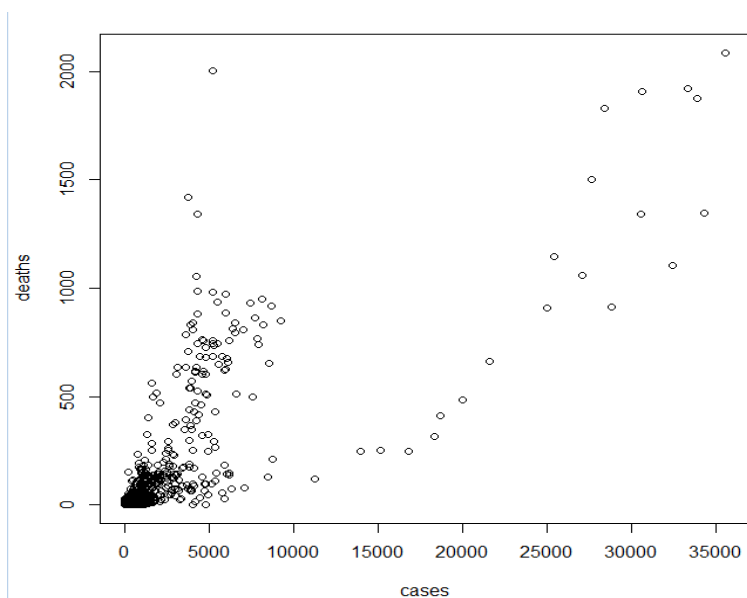


Рисунок 1 – Залежність смертності від захворюваності

Щоб визначити збільшення смертності побудуємо багатофакторну регресійну модель, у якій використовується предиктором буде виступати PopData 2018 (кількість населення). Також будемо враховувати кількість

zareєстрованих випадків Covid-19. На рисунку 2 відображений звіт даної логістичної моделі.

Логістична модель дає змогу передбачити чи буде збільшення смертності. Параметр дисперсії для прийнятої сімейства гаусових рівний 2532,395. Також звернемо увагу що нульове відхилення більше за залишкове приблизно у 3 рази. Також AIC:110692.

Використавши predict() стало зрозуміло що кількість летальних випадків буде тільки збільшуватися. Також щоб дізнатися чи буде збільшуватися кількість летальних випадків, по дням конвертуємо ці ймовірності в up и down якщо буде хоча б 1 смерть в день ми отримаємо up (рис. 3).

```
> summary(glm.fit)

Call:
glm(formula = deaths ~ cases + popData2018, data = A1)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-663.28   -2.40   -2.19   -1.87  1722.46

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.235e+00  5.203e-01   4.295 1.76e-05 ***
cases        5.350e-02  3.756e-04 142.442 < 2e-16 ***
popData2018 -9.624e-09  2.524e-09  -3.812 0.000138 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 2532.395)

Null deviance: 77902255  on 10368  degrees of freedom
Residual deviance: 26250807  on 10366  degrees of freedom
(168 observations deleted due to missingness)
AIC: 110692

Number of Fisher Scoring iterations: 2
```

Рисунок 2 – Звіт підгонки логістичної моделі

```
> glm.pred = rep("Down", 10537)
> glm.pred[glm.probs>1]="Up"
> table(glm.pred,deaths)
      deaths
glm.pred  0    1    2    3    4
Down    413  38  19    9   14
Up     7996 555 279 157 122
```

Рисунок 3 – Прогноз смертельних випадків

Зі звіту можемо побачити тенденцію зростання смертності. Також було вирішено перевірити точність даного передбачення і вона стала рівна 80% це доводить, що передбачення більш ніж правдиві.

Отже, побудована логістична модель визначила залежність ймовірності смертності і показала що погіршення ситуації невідворотне. Автоматично, якщо

збільшується смертність, то захворюваність збільшуватися також. Зв'язок смертності з випадками захворюваності і населеності певних територій показує, що чим більше населеність тим більше смертність. Також за допомогою вбудованого функціоналу пакету R був побудований графік, що яскраво ілюструє дану залежність.

Список літератури

Офіційний сайт Минфину: <https://index.minfin.com.ua/reference/coronavirus/>

Джеймс Г., Уиттон Д., Тибишани Р. Введение в статистическое обучение с примерами R. Изд. Второе, испр. Пер с англ. С.Э. Мاستицкого – М. ДМК Пресс, 2017. -456с.

УДК 004.8:796

*Кадацький М. А., студент 4-го курсу спеціальності 124 «Системний аналіз»
Мельников О. Ю., к. т. н., доцент, доцент кафедри інтелектуальних систем прийняття рішень*

ПОСТАНОВКА ЗАДАЧІ РОЗРАХУНКУ ПОКАЗНИКІВ СПОРТСМЕНА-МЕТАЛЬНИКА ЯДРА ІЗ ЗАСТОСУВАННЯМ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ З 14 ВХІДНИМИ ФАКТОРАМИ

Донбаська державна машинобудівна академія, м. Краматорськ

Сучасний рівень розвитку легкої атлетики, зокрема штовхання ядра, ставить задачу по розробці нових, більш раціональних засобів і методів спортивної підготовки, які сприяють швидкому і надійному досягненню високих спортивних результатів. Але силу безмежно збільшувати не можна, і подальше зростання результатів можливе не стільки за рахунок вдосконалення техніки метань, для чого доцільно використовувати інформаційні технології.

Для проведення розрахунків дальності польоту ядра залежно від початкової швидкості його виштовхування, кута до обрію та висоти над землею, на якій ядро залишає руку, по формулах з [1] була створена інформаційна система – застосунок [2], що дозволяє провести моделювання штовхання ядра з місця та визначити оптимальне сполучення показників для певного ядра. Однак очевидно, що опис спортивної техніки винятково рівняннями механіки може не враховувати ряд факторів, які, будучи малозначущими для абсолютних значень результатів, можуть мати серйозний вплив на відносні показники..

У фізичній культурі та спорті нейронні мережі використовуються для аналізу і прогнозування показників фізичної підготовленості спортсменів, а також результатів спортивних змагань [3]. Для наявних даних з [4] було сформульовано задачу прогнозування: за наявними даними про вік, ріст, масу тіла атлета, а також характеристиках польоту ядра визначити дальність цього