

Испит из статистике
септембар 2018.

1. Посматра се кретање цена берзанског индекса $S\&P$ 500. Циљ истраживања је да се на основу података које поседујемо о цени у претходном периоду предвидимо да ли ће наредног дана доћи до скока или пада цене. Случајну величину (у бази `\$market` означена са `Direction`) чију вредност желимо да прогнозирамо можемо кодирати са две вредности: 1 уколико је дошло до скока, 0 уколико је дошло до пада цене. Подаци које још поседујемо су `Lag1` и `Lag5`, дневни принос претходног дана и пре 5 дана. Напомињемо да је дневни принос дат у процентима и представља релативну промену цене у односу на цену претходног дана.

а) Једно од могућег решења проблема је да посматрамо модел

```
glm.fit=glm(Direction~Lag1+Lag5,family=binomial).
```

Добија се

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|----------|------------|---------|----------|
| (Intercept) | 0.0740 | 0.0567 | 1.31 | 0.1917 |
| Lag1 | -0.0702 | 0.0500 | -1.40 | 0.1607 |
| Lag5 | 0.0091 | 0.0494 | 0.18 | 0.8540 |

Написати формулу модела који смо добили. Описати метод који је био коришћен за добијање оцена модела. (3+3)

б) На основу добијеног модела, колика је оцењена вероватноћа да дође до скока цене акције сутра ако знамо да је данашњи принос 0.95 а пре четири дана је био -2.10? (7)

в) Како бисте на основу прогнозиране вероватноће скока закључили да ли ће до скока доћи? Да ли би се ваш одговор променио уколико бисте на располагању имали још тест података (а не само принос данашњег дана)? (3+3)

3. а) Навести неку ситуацију када је природније користити бајесовски приступ у оцењивању а не фреквенционистички. (3)

б) Како се дефинише Бајесов ризик? (3)

в) Дат је узорак из геометријске $\mathcal{G}(p)$ расподеле: 2, 1, 1, 7, 0, 1, 4. Ако претпоставимо да је априорна оцена за p униформна $\mathcal{U}(0, 1)$, наћи Бајесову оцену параметра p , на основу средњеквадратне функције губитака. (7)

4. Нека је X_1, X_2, \dots, X_n прост случајан узорак из нормалне расподеле $\mathcal{N}(\mu, \sigma^2)$.

а) Оценити непознате параметре μ и σ^2 методом максималне веродостојности. Означити оцене са $\hat{\mu}$ и $\hat{\sigma}^2$ (6)

б) Коју расподелу има $\hat{\mu}$? (5)

в) Одредити 90% интервал поверења за μ ако је реализован узорак

0.938, 1.074, 1.536, 1.504, 0.886, 0.884, 1.179, 1.513, 1.315, 0.193.

Како се интерпретира интервал поверења? (5 + 2)

г) За исти узорак одредити емпиријску функцију расподеле. (3)

д) Коришћењем Колмогоров-Смирнов теста, с нивоом значајности $\alpha = 0.05$, тестирати да је узорак из дела в) из нормалне $\mathcal{N}(1, 0.5)$. Да ли сте исти тест могли да користите у случају да вам нису били познати параметри нормалне расподеле? Одговор образложити. (5 + 5)

5. Дат је следећи узорак: 0.467, 0.132, 0.703, 0.1228441, 0.4303643, 0.7185600, 0.1758184, 0.3751692, 0.1062426, 0.226. Тестирајте, са нивоом значајности теста $\alpha = 0.1$, хипотезу да је медијана узорка 0.5. Приликом одабира одговарајућег теста водите рачуна о расподели обележја на популацији из које је узет узорак. (7)

Напомена: Израда задатака траје 150 минута. Максималан број поена на испиту је 70.

Сваки тачно урађен део задатка вреди онолико поена колико је написано. Да би се добили поени није довољно да само коначан резултат буде тачан већ је потребно објаснити поступак решавања, односно образложити одговор. У случају нетачног или недовршеног решења, могуће је деоно бодовање уколико је поступак решавања исправан и образложен. Нечитки одговори неће бити бодовани. Потписане текстове задатака треба предати.