Контрольная 4. Распределения, сходимости.

1. (3)Пусть $\hat{F}(x)$ — эмпирическая функция распределения. Пусть $x, y \in \mathbb{R}$. Найдите ковариацию $Cov(\hat{F}(x), \hat{F}(y))$.

2. (3)Найти оценку максимального правдоподобия параметра θ у распределения с плотностью

$$f_\theta(y) = \theta y^{\theta-1} \quad y \in [0, 1]$$

3. (3)Пусть $\{X_1, \ldots, X_n\}$ — выборка из распределения Бернулли с параметром p. Является ли статистика $\hat{p} = (\overline{X})^2$ несмещённой оценкой параметра p? Состоятельной?

4. (3)Пусть $\{X_1, \ldots, X_{3n}\}$ — выборка объёма 3n из нормального распределения со средним a и единичной дисперсией. Проверить несмещённость и состоятельность следующих оценок параметра a

$$\frac{1}{n}\sum_{i=n+1}^{2n} X_i$$

$$\frac{1}{n}\sum_{i=1}^{n} X_{3i}$$

5. (3)Пусть $\{X_1, \ldots, X_n\}$ — выборка из равномерного распределения на отрезке $[0, \theta]$. Проверить состоятельность и несмещённость следующих оценок параметра θ

$$(n+1)X_{(1)}$$

$$\frac{n+1}{n}X_{(n)}$$

6. (3)Пусть $\{X_1, \ldots, X_n\}$ — выборка из равномерного распределения на отрезке $[0, \theta]$. Является ли оценка $(n+1)X_{(1)}$ асимптотически нормальной оценкой параметра θ?

7. (3+3)Найти методом моментов и методом максимального правдоподобия оценки параметров распределениея Вейбула

$$f_X(x) = \begin{cases} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}e^{-(\frac{x}{\lambda})^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

8. (4)Распределение Кэптейна определяется плотностью

$$f_\theta(y) = \frac{g^{'}(y)}{\sqrt{2pi}}e^{-\frac{(\theta-g(y))^2}{2}}$$

где $g(y)$ — неубывающая дифференцируемая функция. Найти оценку максимального правдоподобия параметра θ.

9. У Саши есть 2 классификатора: получше и похуже. Предположим, что качество одного классификатора = с и не известно. Классификатор тем точнее и лучше, чем больше с. Предположим, что при измерении качества мы наблюдаем значения $y_i = c + \epsilon_i$. ϵ_i - независимые одинаково распределённые нормальные величины с некоторой дисперсией и нулевым математическим ожиданием. Таким образом, y_i - тоже случайные величины с мат. ожиданием с и дисперсией σ^2.

   (a) (3) Постройте доверительный интервал для значения с так, чтобы вероятность того, что с лежит в этом интервале была не менее 0.95. Найдите границы интервалов для данных каждого классификатора.

   (b) (3) Любым известным вам способом (хоть с лекций, хоть с практик, хоть из гугла) проверьте гипотезу о том, что классификатор получше работает точнее, чем тот, что похуже. Можно ли утверждать, что лучший классификатор точнее.

   Измерения классификатора получше: 0.64175111, 0.63247873, 0.63313111, 0.63270667, 0.63184 , 0.64238667, 0.63818667, 0.64000529, 0.63401333, 0.63696317, 0.63300127, 0.63815111, 0.63456127, 0.63844444, 0.64431556, 0.64572444, 0.63088381, 0.63283492, 0.64132952, 0.6414 , 0.631 , 0.63810667, 0.64361651, 0.63152762, 0.64319556, 0.64393778, 0.64226286, 0.63413079, 0.63395556, 0.62351175, 0.63728095, 0.63190349, 0.63716 , 0.63773397, 0.63676381, 0.63008063, 0.63776952, 0.63110952, 0.63800444, 0.63679111

   Измерения классификатора похуже: 0.62203002, 0.62244152, 0.62315778, 0.6215295 , 0.62170169, 0.62271204, 0.62261512, 0.62374008, 0.62170984, 0.62257677, 0.61856799, 0.6150659 , 0.61587726, 0.62742243, 0.6164587 , 0.62721114, 0.62420552, 0.62223767, 0.62851293, 0.61790504, 0.61169198, 0.62410424, 0.62155702, 0.61685259, 0.61949887, 0.62610344, 0.62347114, 0.61816931, 0.62375206, 0.61676952, 0.62067124, 0.6205655 , 0.61914936, 0.62486339, 0.61320572, 0.61730455, 0.62041464, 0.61807048, 0.62044667, 0.61814111