ФКН ВШЭ, 3 курс, 3 модуль

Материалы к экзамену

Вероятностные модели и статистика случайных процессов, весна 2018

Теоретический максимум

1. С использованием теоремы А. Н. Колмогорова продемонстрировать невозможность существования непрерывного случайного процесса с сечениями, являющимися последовательностью независимых случайных величин.
2. Сформулировать и доказать необходимые и достаточные условия непрерывности случайного процесса в среднем квадратичном.
3. С использованием определения стохастически непрерывного процесса доказать, что свойства стохастической непрерывности и независимости сечений случайного процесса (при близких значениях времени) являются несовместными.
4. Сформулировать и доказать утверждение о необходимых и достаточных условиях гауссовости случайного вектора.
5. Сформулировать и доказать теорему о нормальной корреляции в случае пары гауссовских случайных величин.
6. Сформулировать и доказать теорему о нормальной корреляции в случае пары гауссовских случайных векторов.
7. Сформулировать определения и свойства винеровского и гауссовского процессов. Привести примеры гауссовских процессов. Описать полный набор параметров, однозначно определяющих гауссовский процесс, обосновать это описание.
8. Получить явный аналитический вид системы конечномерных распределений процесса Орнштейна-Уленбека.
9. Перечислить классы стационарности случайных процессов, описать связь между ними. Привести примеры процессов, относящихся к каждому классу, но не относящихся к остальным.
10. Сформулировать и доказать свойства (распределение, матожидание и дисперсию) пуассоновского потока событий с интенсивностью $\lambda > 0$ в момент t.
11. Получить явный аналитический вид системы конечномерных распределений пуассоновского процесса.
12. Сформулировать и доказать теорему о вероятности наблюдения заданной последовательности состояний дискретной марковской цепи.
13. Сформулировать и доказать теорему о вероятности перехода дискретной марковской цепи из одного состояния в другое за n шагов.

14. Описать классификацию состояний дискретной марковской цепи.

15. Сформулировать и доказать условия возвратности либо невозвратности случайного блуждания.

16. Сформулировать и доказать утверждение о том, что возвратность или невозвратность состояния дискретной марковской цепи следует из равенства или неравенства бесконечности величины $\sum\limits_{i=1}^{n} p_{ii}^{(n)}$, соответственно.

17. Сформулировать и доказать утверждение о том, что возвратность или невозвратность состояния дискретной марковской цепи равносильна тому, что вероятность f_i события $\{\exists n \in \mathbb{N} : X_n = i\}$, где n – некоторый момент времени, i – рассматриваемое состояние, равняется либо меньше единицы, соответственно.

18. Сформулировать и доказать утверждение о том, что (1) если одно из состояний цепи нулевое, то и все остальные нулевые, (2) если одно из состояний возвратное, то и все остальные возвратные, (3) если одно из состояний периодическое с периодом d, то и все остальные периодические с периодом d.

19. Вывести формулу средней длительности пребывания дискретной марковской цепи в заданном состоянии.

20. Описать вычислительную разностную схему, позволяющую сгенерировать реализацию гауссовского случайного процесса с помощью стохастического интегрирования по броуновскому движению (на примере процесса Орнштейна-Уленбека).

21. Описать вычислительную схему, позволяющую сгенерировать реализацию гауссовского случайного процесса с помощью разложения Холецкого (на примере процесса фрактального броуновского движения).

22. Описать вычислительную схему, позволяющую сгенерировать реализацию однородного пуассоновского случайного процесса.

23. Описать вычислительную схему, позволяющую сгенерировать реализацию неоднородного пуассоновского случайного процесса.

24. Пусть $(h_1, \ldots, h_n)$ – реализация, полученная в результате наблюдений величин h_k из модели $\mathrm{MA}(q)$ в моменты $k = 1, \ldots, n$, и $\overline{h_n} = \frac{1}{n} \sum\limits_{k=1}^{n} h_k$ – временное среднее. Доказать, что стремление величины $\Delta_n^2 = \mathrm{E}\,|\overline{h_n} - \mu|^2$ к нулю при $n \to \infty$ равносильно стремлению к нулю суммы $\frac{1}{n} \sum\limits_{k=1}^{n} R(k)$, где $R(k) = \mathrm{cov}(h_{n+k}, h_n)$. Здесь $\mathrm{E}\, h_n = \mu$.

25. Вывести уравнения Юла-Уолкера для авторегрессионной модели $\mathrm{AR}(p)$.

26. Какую задачу для графической вероятностной модели позволяет решить алгоритм min-sum belief propagation? Опишите этот алгоритм. Необязательно выписывать окончательную формулу через сообщения, достаточно рекуррентного представления и введения соответствующих функций.

27. Какую задачу для графической вероятностной модели позволяет решить алгоритм sum-product belief propagation? Опишите этот алгоритм. Необязательно выписывать окончательную формулу через сообщения, достаточно рекуррентного представления и введения соответствующих функций.

28. Пусть дана простейшая графическая модель для скрытой марковской модели:

$$\begin{array}{ccccc} \mathrm{T} & \bigcirc \rightarrow & \bigcirc \rightarrow & \dots & \bigcirc \rightarrow \\ & \downarrow & \downarrow & \dots & \downarrow \\ \mathrm{X} & \bigcirc & \bigcirc & \dots & \bigcirc \end{array}$$

   - Сформулируйте соотвествующую ей вероятностную модель в общем виде. Какие три возможных задачи можно поставить?
   - Пусть вероятность перехода задается матрицей A, вектор скрытых состояний T дискретный. Рассмотрим ситуацию обучения с учителем, а именно:

$$p(X_{tr}, T_{tr}|\theta) \to \max_{\theta} \tag{1}$$

     Получите оценку для A.

29. Пусть дана простейшая графическая модель для скрытой марковской модели:

$$\begin{array}{ccccc} \mathrm{T} & \bigcirc \rightarrow & \bigcirc \rightarrow & \dots & \bigcirc \rightarrow \\ & \downarrow & \downarrow & \dots & \downarrow \\ \mathrm{X} & \bigcirc & \bigcirc & \dots & \bigcirc \end{array}$$

   - Сформулируйте соотвествующую ей вероятностную модель в общем виде. Какие три возможных задачи можно поставить?
   - Рассмотрите задачу при неизвестных скрытых состояниях T. Опишите ее решение. Достаточно получить функционал и дать описание, как применять оптимизационные методы min-sum и min-product.

30. Сформулировать задачу принятия решения по фиксированному числу наблюдений. Определить рандомизированный и нерандомизированный тесты, ошибки первого и второго родов. Сформулировать лемму Неймана-Пирсона и условия ее оптимальности.

31. Сформулировать последовательную задачу принятия решения. Определить среднее время принятия решения, ошибки первого и второго родов. Сформулировать последовательный критерий отношения правдоподобия и условия его оптимальности.

32. Сформулировать задачу о разладке случайной последовательности. Записать критерий Лордена в задаче о разладке. Записать статистику кумулятивных сумм для независимых наблюдений в этой задаче.

33. Сформулировать задачу о разладке случайной последовательности. Записать критерий Ширяева-Робертса в задаче о разладке. Записать статистику Ширяева-Робертса для независимых наблюдений в этой задаче.

34. Сформулировать задачу о разладке случайной последовательности. Записать статистику контрольных карт Шухарта в этой задаче.

Задачи

1. Подсчитайте математическое ожидание, дисперсию и ковариационную функцию случайного процесса $Y = (Y_t)_{t \geqslant 0}$, задаваемого соотношением

$$Y_t = a(t)X_t + b(t),$$

где $X = (X_t)_{t \geqslant 0}$ – случайный процесс с математическим ожиданием $m(t) = \mathrm{E}\, X_t$, дисперсией $\sigma^2(t) = \mathrm{E}[X_t - \mathrm{E}\, X_t]^2$ и ковариационной функцией $R(t_1, t_2) = \mathrm{E}[(X_{t_1} - \mathrm{E}\, X_{t_1})(X_{t_2} - \mathrm{E}\, X_{t_2})]$.

2. Доказать, что пуассоновский поток событий является стохастически непрерывным случайным процессом.

3. Пусть $\boldsymbol{\xi} = (\xi_1, \xi_2)^\intercal$ – гауссовский случайный вектор с математическим ожиданием $\mathrm{E}\, \boldsymbol{\xi} = \boldsymbol{\mu} = (\mu_1, \mu_2)^\intercal$ и ковариационной матрицей

$$\mathrm{E}[(\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})^\intercal] = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Выписать явное аналитическое выражение для двумерной плотности распределения случайного вектора $\boldsymbol{\xi}$.

4. Пусть $\boldsymbol{\xi} = (\xi_1, \xi_2)^\intercal$ – гауссовский случайный вектор с математическим ожиданием $\mathrm{E}\, \boldsymbol{\xi} = \boldsymbol{\mu} = (\mu_1, \mu_2)^\intercal$ и ковариационной матрицей

$$\mathrm{E}[(\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})^\intercal] = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Подсчитать явное аналитическое выражение для условной плотности $f_{\xi_2|\xi_1}(x_2|x_1)$ распределения случайного вектора ξ_2 при условии $\xi_1 = x_1$.

5. Пусть $N^1, N^2, \ldots, N^n$ – независимые пуассоновские потоки событий с интенсивностями $\lambda_1, \lambda_2, \ldots, \lambda_n$, соответственно. Определить тип и параметры процесса $N_t = \sum\limits_{i=1}^{n} N_t^i$.

6. Подсчитать корреляции процесса MA(2).

7. Подсчитать корреляции процесса MA(q).

8. Подсчитать математическое ожидание процесса AR(1) и его предел при $n \to \infty$, где n – время.

9. Подсчитать дисперсию процесса AR(1) и ее предел при $n \to \infty$, где n – время.

10. Подсчитать ковариацию процесса AR(1) и ее предел при $n \to \infty$, где n – время.

11. Записать правдоподобие $L(h_1, \ldots, h_n|\boldsymbol{\theta})$ выборки $(h_1, \ldots, h_n)$ из авторегрессионной модели AR(p), где $\boldsymbol{\theta} = (a_0, \ldots, a_p, \sigma)$.

12. Подсчитать математическое ожидание процесса ARMA(1,1) и его предел при $n \to \infty$, где n – время.

13. Подсчитать дисперсию процесса ARMA(1,1) и е предел при $n \to \infty$, где n – время.

14. Подсчитать ковариацию процесса ARMA(1,1) и ее предел при $n \to \infty$, где n – время.

15. Подсчитать математическое ожидание квадрата процесса ARCH(1) (величину $\mathrm{E}\, h_n^2$) и его предел при $n \to \infty$, где n – время.

16. Подсчитать математическое ожидание 4 степени процесса ARCH(1) (величину $\mathrm{E}\, h_n^4$) и его предел при $n \to \infty$, где n – время.

17. Подсчитать дисперсию квадрата процесса ARCH(1) (величину $\mathrm{D}h_n^2$) и ее предел при $n \to \infty$, где n – время.

18. Подсчитать первую корреляцию процесса ARCH(1) (величину $\rho_1 = \mathrm{E}\, h_n^2 h_{n-1}^2$) и ее предел при $n \to \infty$, где n – время.

19. Записать правдоподобие $L(h_1, \ldots, h_n | \boldsymbol{\theta})$ выборки $(h_1, \ldots, h_n)$ из модели ARCH(1), где $\boldsymbol{\theta} = (\alpha_0, \alpha_1)$.

20. Подсчитать математическое ожидание квадрата процесса GARCH(1,1) (величину $\mathrm{E}\, h_n^2$) и его предел при $n \to \infty$, где n – время.

21. Подсчитать математическое ожидание 4 степени процесса GARCH(1,1) (величину $\mathrm{E}\, h_n^4$) и его предел при $n \to \infty$, где n – время.

22. Подсчитать первую корреляцию процесса GARCH(1,1) (величину $\rho_1 = \mathrm{E}\, h_n^2 h_{n-1}^2$) и ее предел при $n \to \infty$, где n – время.

23. Записать правдоподобие $L(h_1, \ldots, h_n | \boldsymbol{\theta})$ выборки $(h_1, \ldots, h_n)$ из модели GARCH(1,1), где $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1)$.

24. Приведите пример некоррелированных, но зависимых случайных величин.

25. Приведите пример процесса, являющегося сильно стационарным, но не эргодичным по математическому ожиданию в среднем квадратичном.

26. Рассмотрим марковскую цепь, изображённую на рисунке 1. На ней присутствуют 2 рекуррентных класса: $R_1 = 1, 2$, $R_2 = 5, 6, 7$. Пусть $X0 = 3$. Найти вероятность того, что цепь будет поглощена в R_1.

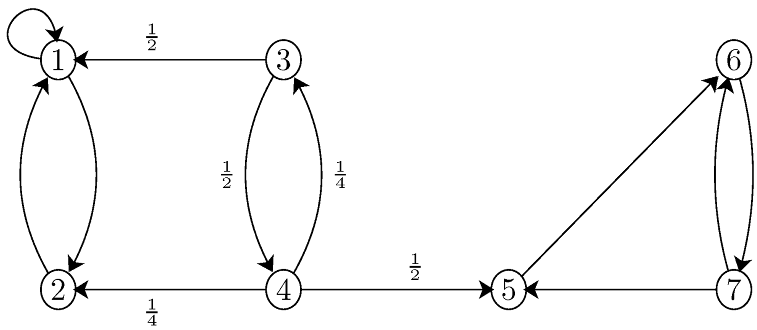

Рис. 1: Рисунок к задаче 26.

27. Дана матрица перехода
$$P = \begin{pmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{pmatrix}.$$
Найти P^n.

28. Игрок вступает в игру с капиталом 100\$. В каждом ходу игры игрок получает 1\$ с вероятностью p и теряет 1\$ с вероятностью $1-p$. Игра продолжается, пока игрок не наберёт 300\$ или не проиграет все деньги. Какова вероятность, что игра когда-нибудь закончится? Какова вероятность, что игрок выйдет победителем?

29. Рассмотрим марковскую цепь, показанную на рисунке 2. Положим $\frac{1}{2} < p < 1$. Есть ли у данной цепи предельное распределение? Найти
$$\lim_{n\to\infty} \mathrm{P}(X_n = j | X_0 = i).$$

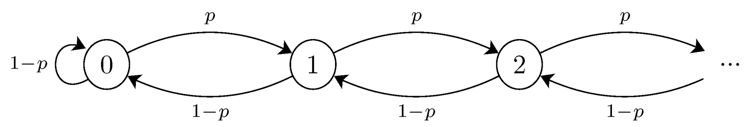

Рис. 2: Рисунок к задаче 29.

30. Доказать, что функция $R(t,s) = \min\{t,s\} - ts$ может или не может являться ковариационной функцией случайного процесса.

31. Доказать, что функция $R(t,s) = \min\{t,s\} - t(s+1)$ может или не может являться ковариационной функцией случайного процесса.

32. Пусть $B = (B_t)_{t\geqslant 0}$ – винеровский процесс. Доказать, что процесс
$$C_t = \begin{cases} 0, & t = 0, \\ tB_{1/t}, & t > 0. \end{cases}$$
также винеровский.

33. Пусть $B = (B_t)_{t\geqslant 0}$ – винеровский процесс. Доказать, что процесс
$$C_t = \sqrt{c}B_{t/c}, \qquad c = \mathrm{const} > 0$$
также винеровский.

34. $\xi_1, \ldots, \xi_n$ – независимые одинаково распределенные показательные случайные величины. Подсчитать (по индукции) плотность распределения суммы $\xi_1 + \ldots + \xi_n$.

35. Пусть $N = (N_t)_{t\geqslant 0}$ – пуассоновский случайный процесс с параметром λ. Доказать, что случайный процесс $M = (M_t)_{t\geqslant 0}$, задаваемый соотношением $M_t = N_{t+1} - N_t$, является стационарным второго порядка процессом, т.е. что его математическое ожидание $\mathrm{E}\, M_t$ не зависит от времени, а его ковариационная функция $R_M(t_1,t_2)$ зависит от t_1 и t_2 через их разность $\tau = t_1 - t_2$.

36. Доказать положительную определенность функции

$$R(t,s) = \begin{cases} 1 - |t-s|, & |t-s| < 1, \\ 0, & |t-s| >= 1. \end{cases}$$

37. Доказать положительную определенность функции

$$R(t,s) = e^{-|t-s|}.$$

38. Пусть $\xi_1, \xi_2, \dots$ – независимые одинаково распределенные случайные величины, причем $\mathrm{P}(\xi_t = 1) = p, \mathrm{P}(\xi_t = -1) = 1 - p$. Является ли цепью Маркова последовательность $\eta_t = \xi_t \xi_{t+1}$? Если да, найти ее вероятности перехода за один шаг.

39. Пусть $\xi_1, \xi_2, \dots$ – независимые одинаково распределенные случайные величины, причем $\mathrm{P}(\xi_t = 1) = p, \mathrm{P}(\xi_t = -1) = 1 - p$. Является ли цепью Маркова последовательность $\eta_t = \xi_1 \xi_2 \dots \xi_t$? Если да, найти ее вероятности перехода за один шаг.

40. Для случайного процесса $h = (h_t)_{t \geqslant 0}$, заданного выражением

$$h_t = \varepsilon_t - 1.3\varepsilon_{t-1} + 0.4\varepsilon_{t-2},$$

где $\varepsilon = (\varepsilon_t)_{t \geqslant 0}$ – последовательность независимых одинаково стандартно нормально распределенных случайных величин,

(a) Провести исследование на стационарность 2-го порядка (с доказательством);

(b) Вычислить первые четыре автокорреляции $\rho_0, \rho_1, \rho_2, \rho_3$;

(c) Вычислить дисперсию случайной величины h_t.

41. Для случайного процесса $h = (h_t)_{t \geqslant 0}$, заданного выражением

$$h_t - h_{t-1} = \varepsilon_t - 1.3\varepsilon_{t-1} + 0.3\varepsilon_{t-2},$$

где $\varepsilon = (\varepsilon_t)_{t \geqslant 0}$ – последовательность независимых одинаково стандартно нормально распределенных случайных величин,

(a) Провести исследование на стационарность 2-го порядка (с доказательством);

(b) Вычислить первые четыре автокорреляции $\rho_0, \rho_1, \rho_2, \rho_3$;

(c) Вычислить дисперсию случайной величины h_t.

42. Для модели GARCH(1, 1) временного ряда, задающейся уравнениями

$$X_n = \mu + h_n, \quad h_n = \sigma_n \varepsilon_n,$$
$$\sigma_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2 + \beta_1 \sigma_{n-1}^2,$$

где $\varepsilon = (\varepsilon_t)_{t=1,2,\dots}$ – процесс гауссовского белого шума,

(a) Записать формулу для подсчета σ_{n+1}^2;

(b) Подсчитать распределение величины X_{n+1}.

43. По выборке $(X_1, \dots, X_n)$ из биномиального распределения $\mathrm{Bin}(k, p)$ построить критерий Неймана-Пирсона для проверки гипотезы $\mathbb{H}_0 : p = p_0$ против альтернативы $\mathbb{H}_1 : p = p_1$, где $0 < p_0 < p_1 < 1$.

44. Дана выборка $(X_1, \dots, X_n)$ из нормального $\mathcal{N}(\mu, \sigma^2)$ распределения. Построить критерий проверки гипотезы $\mathbb{H}_0 : \mu = 0$ против альтернативы $\mathbb{H}_1 : \mu = a$ и определить наименьший объем выборки, при котором вероятности ошибок 1 и 2 родов не превышают заданного значения δ.

45. По выборке $(X_1, \dots, X_n)$ из пуассоновского распределения $\Pi(\lambda)$ построить критерий Неймана-Пирсона для проверки гипотезы $\mathbb{H}_0 : \lambda = \lambda_0$ против альтернативы $\mathbb{H}_1 : \lambda = \lambda_1$, где $0 < \lambda_0 < \lambda_1$.

46. В последовательности $\xi_1, \dots, \xi_n$ независимых испытаний, выполненных согласно схеме Бернулли, $\mathrm{P}(\xi_i = 1) = p, \mathrm{P}(\xi_i = 0) = 1-p$. Построить критерий проверки гипотезы $\mathbb{H}_0 : p = p_0$ против альтернативы $\mathbb{H}_1 : p = p_1$ и определить наименьший объем выборки, при котором вероятности ошибок 1 и 2 родов не превышают заданного значения δ.

47. Основная гипотеза состоит в том, что данный человек лишён телепатических способностей и угадывает мысли на расстоянии в каждом единичном эксперименте с вероятностью $\frac{1}{2}$. Гипотеза же о наличии телепатических способностей у данного человека принимается, если в 100 независимых однотипных экспериментах по угадыванию мыслей на расстоянии не менее 70 заканчиваются успехом. Чему равна вероятность признать телепатом человека без телепатических способностей?

48. Доход от проданной газеты равен A ($=$ розничная цена – оптовая), потери от непроданной равны B (оптовая цена). Число покупателей, приходящих в киоск в день, моделируется сл.вел. X с функцией распределения $F(x)$. Для ее оценки можно использовать записи прошлых продаж. Сколько газет следует брать для продажи ?

49. Путем выборочного опроса проверяется гипотеза о том, что стиральным порошком фирмы A пользуется 30% населения против гипотезы, что им пользуется только 20% населения. Оцените объем выборки, необходимый для проверки гипотезы с ошибкой первого рода не более 5% и второго рода не более 2.5%.

50. Пусть $X_1, ..., X_n$ – простая выборка из нормального распределения со средним a и дисперсией 1. Для проверки основной гипотезы $a = 0$ против альтернативы $a = 1$ используется следующий критерий: основная гипотеза принимается, если $X_{(n)} = \max\limits_{i=1,\dots,n} X_i < 3$ и отвергается в противном случае. Найти вероятности ошибок первого и второго родов.

51. В задаче обнаружения разладки наблюдается последовательность $X_1, \dots, X_k, \dots$ независимых нормально $\mathcal{N}(\mu, 1)$ распределенных случайных величин, причем до момента разладки $\mu = 0$, а после момента разладки $\mu = m$. Построить процедуру обнаружения разладки на основе статистики кумулятивных сумм. Привести формулы для отношения правдоподобия, а также для одного шага итеративного алгоритма кумулятивных сумм. В какой момент следует поднимать тревогу об обнаружении разладки?

52. В задаче обнаружения разладки наблюдается последовательность $X_1, \ldots, X_k, \ldots$ независимых нормально $\mathcal{N}(0, \sigma^2)$ распределенных случайных величин, причем до момента разладки $\sigma^2 = 1$, а после момента разладки $\sigma^2 = s^2$. Построить процедуру обнаружения разладки на основе статистики кумулятивных сумм. Привести формулы для отношения правдоподобия, а также для одного шага итеративного алгоритма кумулятивных сумм. В какой момент следует поднимать тревогу об обнаружении разладки?

53. В задаче обнаружения разладки наблюдается последовательность $X_1, \ldots, X_k, \ldots$ независимых одинаково распределенных случайных величин, причем до момента разладки X_i — нормально $\mathcal{N}(1, 1)$ распределенные случайные величины, а после момента разладки X_i — распределены экспоненциально $\mathrm{Exp}(1)$. Построить процедуру обнаружения разладки на основе статистики Ширяева-Робертса. Привести формулы для отношения правдоподобия, а также для одного шага итеративного алгоритма Ширяева-Робертса. В какой момент следует поднимать тревогу об обнаружении разладки?

54. В задаче обнаружения разладки наблюдается последовательность $X_1, \ldots, X_k, \ldots$ независимых экспоненциально $\mathrm{Exp}(\lambda)$ распределенных случайных величин, причем до момента разладки $\lambda = 1$, а после момента разладки $\lambda = m$. Построить процедуру обнаружения разладки на основе статистики Ширяева-Робертса. Привести формулы для отношения правдоподобия, а также для одного шага итеративного алгоритма Ширяева-Робертса. В какой момент следует поднимать тревогу об обнаружении разладки?

55. В задаче обнаружения разладки наблюдается последовательность $X_1, \ldots, X_k, \ldots$ независимых бернуллиевских случайных величин, причем до момента разладки $p = 1/2$, а после момента разладки $p \notin [1/2 - \varepsilon, 1/2 + \varepsilon]$. Построить процедуру обнаружения разладки на основе статистики контрольных карт. Привести формулы для отношения правдоподобия, а также для одного шага алгоритма контрольных карт. В какой момент следует поднимать тревогу об обнаружении разладки?