**ЗАДАНИЕ**

на практическое занятие № 4 по дисциплине

«Прикладная теория вероятностей и статистика»

I. Тема: Реализация на ЭВМ алгоритмов оценивания среднего риска.

II. Исходные данные:

1. Производится контроль температурного режима на борту космического аппарата (КА). Возможны два технических состояния:

а) КА находится в исправном состоянии *S*0 (гипотеза *Н*0);

б) КА находится в неисправном состоянии *S*1 (гипотеза *Н*1).

2. При превышении граничного значения температуры *х*гр техническое состояние КА считается неисправным и КА не способен выполнять целевую задачу.

3. Для исправного состояния *S*0 среднее значение температуры составляет и среднее квадратическое отклонение . При неисправном состоянии *S*1 параметры равны и . Плотности вероятностей распределения температур на борту КА предполагаются нормальными.

4. По статистическим данным априорные вероятности исправного и неисправного состояний КА соответственно равны и

5. Значения потерь принять равными:

, (*N*+5)/*N* усл. ед., усл. ед.

Таблица 1. Варианты исходных данных

| № варианта | Значение параметра | | | | № варианта | Значение параметра | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
| 1 | 18 | 1,8 | 30 | 3,0 | 16 | 19 | 3,8 | 31 | 6,2 |
| 2 | 19 | 1,9 | 31 | 3,1 | 17 | 20 | 1,9 | 32 | 3 |
| 3 | 20 | 2,0 | 32 | 3,2 | 18 | 21 | 2,0 | 33 | 4,8 |
| 4 | 21 | 2,1 | 33 | 3,3 | 19 | 19 | 1,9 | 33 | 3,3 |
| 5 | 22 | 2,2 | 34 | 3,4 | 20 | 20 | 2,0 | 34 | 3,4 |
| 6 | 23 | 2,3 | 35 | 3,5 | 21 | 21 | 2,1 | 35 | 3,5 |
| 7 | 24 | 2,4 | 36 | 3,6 | 22 | 22 | 2,2 | 36 | 3,6 |
| 8 | 18 | 2,7 | 30 | 4,5 | 23 | 23 | 1,5 | 30 | 3,0 |
| 9 | 19 | 2,9 | 31 | 4,7 | 24 | 24 | 1,4 | 31 | 3,7 |
| 10 | 20 | 3,0 | 32 | 4,8 | 25 | 18 | 1,7 | 32 | 3,8 |
| 11 | 21 | 3,2 | 33 | 5,0 | 26 | 19 | 1,9 | 33 | 4,0 |
| 12 | 22 | 3,3 | 34 | 5,1 | 27 | 20 | 2,0 | 34 | 4,1 |
| 13 | 23 | 3,5 | 35 | 5,2 | 28 | 21 | 1,2 | 35 | 4,2 |
| 14 | 24 | 3,6 | 36 | 5,4 | 29 | 22 | 1,3 | 36 | 4,4 |
| 15 | 18 | 3,6 | 30 | 6,0 | 30 | 23 | 1,5 | 30 | 4,0 |

III. Выполнить:

1. Заполнить таблицу 2 в соответствии с вариантом *N* при известных значениях параметров α и β

, .

Таблица 2. Параметры состояния КА при известных значениях параметров α и β

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Состояние | *S*0 | | *S*1 | |
| Априорная  вероятность | (КА исправен) | | (КА неисправен) | |
| Принятое  решение | *Н*00 | *Н*10 | *Н*01 | *Н*11 |
| Оценка  решения | правильное | ошибка 1-го рода | ошибка 2-го рода | правильное |
| Вероятность решения | 1 – α | α | β | 1 – β |
| Стоимость  потерь | *C*00 | *C*10 | *C*01 | *C*11 |
| Средний  риск |  |  |  |  |

1.1. Определить вероятность ошибочного решения оператора при известных априорных вероятностях и

1.2. Определить средний риск (потери), связанный с ошибочными решениями оператора, при условии, что известны априорные вероятности и состояния КА и матрица потерь.

2. Определить граничное значение температуры *х*гр, выше которого техническое состояние КА считается неисправным, а также вероятность ложной тревоги *P*(*H*01), вероятность пропуска дефекта *P*(*H*10) и средний риск *R*:

а) по методу минимального среднего риска;

б) по методу минимальной вероятности ошибочного решения.

3. Результаты выполнения задания представить в виде таблицы 3.

Таблица 3. Определение граничного значения температуры *х*гр и среднего риска *R*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | Метод | Граничное значение | Вероятность ложной  тревоги | Вероятность пропуска  дефекта | Средний риск |
| 1 | Метод минимального среднего риска |  |  |  |  |
| 2 | Метод минимальной вероятности ошибочного решения |  |  |  |  |

4. Сделать выводы по работе.

Номер варианта *N* соответствует порядковому номеру в списке учебной группы.

**Краткие теоретические сведения**

Вводится количественный показатель точности оценивания – *функция риска*, или, по-другому, *функция потерь*.

🞿 ***Определение 4.1.***Функция риска  есть количественная мера  
соответствия принимаемой оценки  истинному значению параметра *х*. ▲

Одновременно функция риска  служит и мерой качества алгоритма , дающего оценку :

. (4.1)

Функция риска  является функцией двух аргументов – принимаемой оценки  (алгоритма , дающего эту оценку ) и истинного значения параметра *х*.

Функция риска  должна быть такой, чтобы отражать следующее интуитивное положение: чем больше отличие величины  от *х*, тем больше должно быть значение , то есть тем больше должен быть риск или больше должны быть возможные потери.

🟆 ***Пример 4.1.*** Для непрерывных величин функция риска  – это поверхность в трехмерном пространстве (рис. 4.1). ▲

Рисунок иллюстрирует, что при совпадении значений оценки  и истинного значения *х* потерь нет, и функция риска  вырождается в линию  = *х*.









*Рис. 4.1. График функции риска для непрерывных величин*

🟆 ***Пример 4.2.*** Для дискретных величин каждой паре возможных значений , , можно поставить в соответствие некоторое число , характеризующее потери (риски), возникающие при принятии в качестве оценки значения  при истинном значении . Все такие числа, а их *N*2, сведем в квадратную матрицу – так называемую *матрицу потерь*:

. ▲

***Замечание 4.1.*** Для непрерывных величин достаточно задавать функцию потерь в виде функции одного аргумента: . ▲

***Замечание 4.2.*** При решении баллистических задач разность  между оценкой  и истинным значением параметра *х* называется *невязкой*. ▲

Пусть производится контроль технического состояния системы терморегулирования (СТР) космического аппарата (КА) по величине температуры (параметр *x*).

**Задача состоит в выборе значения** параметра *x* таким образом, что при следует принимать решение о неисправности СТР (гипотеза ), а при допускать дальнейшую работу (принимается гипотеза ).

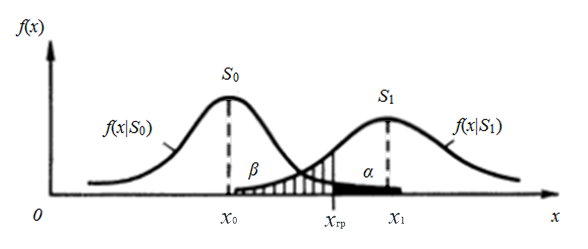
Так как состояние системы характеризуется одним параметром, то система имеет одномерное пространство признаков. Разделение производится на два класса. Условимся считать: – исправное состояние, – наличие дефекта. Тогда правило решения состоит в следующем:

|  |  |
| --- | --- |
| при принимается гипотеза ();  при принимается гипотеза (). | (4.2) |

Температура неоднозначно характеризует состояние СТР. В зависимости от ряда факторов распределение для дефектных и исправных СТР показано на рис. 4.2.

Существенно, что области исправного и дефектного состояний пересекаются и принципиально невозможно выбрать значение , при котором правило (4.2) не давало бы ошибочных решений.

Задача состоит в том, чтобы выбор был в некотором смысле оптимальным, например, давал наименьшее число ошибочных решений.



*Рис. 4.2. Статистические распределения плотности вероятности*

*диагностического параметра для исправного и дефектного состояний.*

Рассмотрим возможные ошибки при принятии решения.

**Ошибкой первого рода α (*ложной тревогой*)** называется случай, когда принимается решение о наличии дефекта, но в действительности система находится в исправном состоянии (вместо принимается ).

**Ошибка второго рода β (*пропуск цели (дефекта)*)** – принятие решения об исправном состоянии, тогда как система содержит дефект (вместо принимается ).

В теории контроля эти ошибки называются риском поставщика и риском заказчика. Очевидно, что эти ошибки могут иметь различные последствия или различные цены.

Обозначим () возможные решения по правилу (4.1) (первый нижний индекс соответствует индексу принятого диагноза, второй – индексу действительного состояния). Тогда – пропуск дефекта, – ложная тревога; и – правильные решения.

Рассмотрим вероятность ложной тревоги при использовании правила (4.2) (случай, когда при объект является исправным, но по правилу (4.2) рассматривается как дефектный). Площадь под кривой плотности вероятности исправного состояния при , выражает условную вероятность ситуации для исправных изделий

|  |  |
| --- | --- |
|  | (4.3) |

Вероятность ложной тревоги равна вероятности произведения двух событий: наличия исправного состояния и значения . Тогда

|  |  |
| --- | --- |
|  | (4.4) |

где – априорная вероятность диагноза (считается известной на основании предварительных статистических данных).

Подобным образом находится вероятность пропуска дефекта

|  |  |
| --- | --- |
|  | (4.5) |

Вероятность принятия ошибочного решения слагается из вероятностей ложной тревоги и пропуска дефекта. Если приписать «цены» этим ошибкам, то получим выражение для среднего риска

|  |  |
| --- | --- |
|  | (4.6) |

Разумеется, цена ошибки имеет условное значение, но она должна учесть предполагаемые последствия ложной тревоги и пропуска дефекта.

В задачах надежности стоимость пропуска дефекта обычно существенно больше стоимости ложной тревоги (). Иногда вводится цена правильных решений и , которая для сравнения со стоимостью потерь (ошибок) принимается отрицательной.

В общем случае средний риск (ожидаемая величина потери) выражается равенством

|  |  |
| --- | --- |
|  | (4.7) |

Величина , предъявляемая для распознавания, является случайной, и поэтому равенства (4.6) и (4.7) представляют собой среднее значение (математическое ожидание) риска.

**Метод минимального среднего риска.**

Найдем граничное значение в правиле (4.2) из условия минимума среднего риска. Дифференцируя (4.7) по и приравнивая производную нулю, получим сначала условие экстремума

|  |  |
| --- | --- |
|  | (4.8) |

или

|  |  |
| --- | --- |
|  | (4.9) |

Это условие часто определяет два значения , из которых одно соответствует минимуму, второе – максимуму риска.

Соотношение (4.9) является необходимым, но недостаточным условием минимума. Для существования минимума в точке вторая производная должна быть положительной , что приводит к следующему условию относительно производных плотностей распределений:

|  |  |
| --- | --- |
|  | (4.10) |

Если распределения и являются, как обычно, одномодальными (т.е. содержат не более одной точки максимума), то при

|  |  |
| --- | --- |
|  | (4.11) |

условие (4.10) выполняется.

Действительно, в правой части равенства стоит положительная величина, а при производная , тогда как при значение .

Для полимодальных распределений условие (4.10) должно проверяться в каждой точке экстремума.

Из условия (4.9) следует, что применительно к параметру *х* решение об отнесении объекта (КА) к состоянию или можно связать с величиной отношения правдоподобия. Напомним, что **отношение плотностей вероятностей распределения при двух состояниях называется отношением правдоподобия**.

По методу минимального среднего риска принимается следующее решение о состоянии объекта, имеющего данное значение параметра :

|  |  |
| --- | --- |
| , если ; | (4.12) |
| , если | (4.13) |

Эти условия вытекают из соотношений (4.10) и (4.9).

Условие (4.12) соответствует , условие (4.13) . Величина представляет собой пороговое значение для отношения правдоподобия. Напомним, что диагноз соответствует исправному состоянию, – дефектному состоянию объекта; – цена ложной тревоги; – цена пропуска цели (первый индекс – принятое состояние, второй – действительное); , – цены правильных решений (условные выигрыши).

В большинстве практических задач условные выигрыши (поощрения) для правильных решений не вводятся и тогда

|  |  |
| --- | --- |
|  | (4.14) |

Часто оказывается удобным рассматривать не отношение правдоподобия, а логарифм этого отношения. Это не изменяет результата, так как логарифмическая функция возрастает монотонно вместе со своим аргументом.

Расчет для нормального и некоторых других распределений при использовании **логарифма отношения правдоподобия** оказывается несколько проще.

Рассмотрим случай, когда параметр имеет нормальное распределение при исправном и неисправном состояниях.

В рассматриваемом случае плотности распределений

Внося эти соотношения в равенство (6.9), получаем после логарифмирования

Из этого уравнения

определяется граничное значение в правиле (4.2) из условия минимума среднего риска.

**Метод минимальной вероятности ошибки.**

Вероятность ошибки (ошибочного решения) для решающего правила (4.2)

|  |  |
| --- | --- |
|  | (4.15) |

Из условия экстремума этой вероятности получаем

|  |  |
| --- | --- |
|  | (4.16) |

Условие минимума дает

|  |  |
| --- | --- |
|  | (4.17) |

или

|  |  |
| --- | --- |
|  | (4.18) |

Как указывалось, для одномодальных распределений при условии (4.11) неравенство (4.15) выполняется и минимум вероятности ошибочного решения получается из соотношения (4.13):

|  |  |
| --- | --- |
|  | (4.19) |

где, как и раньше, , – априорные вероятности диагнозов.

Решение принимается при

|  |  |
| --- | --- |
|  | (4.20) |

и принимается при

|  |  |
| --- | --- |
|  | (4.21) |

Очевидно, что соотношения (4.19)-(4.21) являются частным случаем условия минимального среднего риска, если стоимости решений одинаковы.

Условие выбора граничного значения (4.19) часто называется *условием Зигерта – Котельникова* (условием идеального наблюдателя).

После логарифмирования (4.21) получаем уравнение

из которого определяется граничное значение в правиле (4.2) из условия минимума ошибочного решения оператора.

В задачах надежности рассматриваемый метод часто дает «неосторожные решения», так как последствия ошибочных решений существенно различаются между собой. Обычно цена пропуска дефекта существенно выше цены ложной тревоги. Если указанные стоимости приблизительно одинаковы (для дефектов с ограниченными последствиями, для некоторых задач контроля и др.), то применение метода вполне оправдано.

**Порядок выполнения работы**

1. Заполнить таблицу 2 в соответствии с вариантом *N* при известных значениях параметров α и β

, .

Таблица 2. Параметры состояния КА при известных значениях параметров α и β

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Состояние | *S*0 | | *S*1 | |
| Априорная  вероятность | (КА исправен) | | (КА неисправен) | |
| Принятое  решение | *Н*00 | *Н*10 | *Н*01 | *Н*11 |
| Оценка  решения | правильное | ошибка 1-го рода | ошибка 2-го рода | правильное |
| Вероятность решения | 1 – α | α | β | 1 – β |
| Стоимость  потерь | *C*00 | *C*10 | *C*01 | *C*11 |
| Средний  риск |  |  |  |  |

1.1. Определить вероятность ошибочного решения оператора при известных априорных вероятностях и

1.2. Определить средний риск (потери), связанный с ошибочными решениями оператора, при условии, что известны априорные вероятности и состояния КА и матрица потерь

2. Определить граничное значение температуры *x*0, выше которого техническое состояние КА считается неисправным, вероятность ложной тревоги α, вероятность пропуска дефекта β и средний риск *R* по методу минимального среднего риска;

2.1 Параметр имеет нормальное распределение при исправном и неисправном состояниях.

В рассматриваемом случае плотности распределений

Внося эти соотношения в равенство (6.9), получаем после логарифмирования

2.2. После приведения данного уравнения к стандартному виду *ах*2+*bх*+*c* = 0 определяется граничное положительное значение в правиле (4.2) из условия минимума среднего риска.

*х*гр = (–*b*±(*b*2–4*ac*)^0,5)/2*a*.

2.3. Определяется вероятность ложной тревоги *P*(*H*01) = *P*(*S*0)·α.

Вероятность ложной тревоги равна вероятности произведения двух событий: наличие исправного состояния и значение . Тогда

|  |  |
| --- | --- |
|  |  |

где – априорная вероятность диагноза (считается известной на основании предварительных статистических данных).

Для вычисления вероятности можно использовать формулу Excel «НОРМ.РАСП(*x*гр; среднее значение; СКО; 0 или 1), где параметры выглядят следующим образом: 0 (ЛОЖЬ) –для плотности распределения; 1 (ИСТИНА) – для функции распределения.

2.4. Определяется вероятность пропуска дефекта *P*(*H*10) = *P*(*S*1)·β.

Вероятность пропуска дефекта равна вероятности произведения двух событий: наличие неисправного состояния и значения . Тогда

2.5. Определяется средний риск *R*

3. Определить граничное значение температуры *x*гр, выше которого техническое состояние КА считается неисправным, вероятность ложной тревоги α, вероятность пропуска дефекта β и средний риск *R* по методу минимальной вероятности ошибочного решения (не учитываются функции потерь).

3.1 Параметр имеет нормальное распределение при исправном и неисправном состояниях.

С учетом равенства (4.9) после логарифмирования получаем

3.2. После приведения данного уравнения к стандартному виду *ах*2+*bх*+*c* = 0 определяется граничное положительное значение в правиле (4.2) из условия минимума вероятности ошибки

*х*гр = (–*b*±(*b*2–4*ac*)^0,5)/2*a*.

3.3. Определяется вероятность ложной тревоги *P*(*H*01) = *P*(*S*0)·α.

Вероятность ложной тревоги равна вероятности произведения двух событий: наличие исправного состояния *S* и значения . Тогда

|  |  |
| --- | --- |
|  |  |

где – априорная вероятность диагноза (считается известной на основании предварительных статистических данных).

Для вычисления вероятности можно использовать формулу Excel «НОРМ.РАСП(*x*гр; среднее значение; СКО; 0 или 1), где параметры выглядят следующим образом: 0 (ЛОЖЬ) –для плотности распределения; 1 (ИСТИНА) – для функции распределения.

3.4. Определяется вероятность пропуска дефекта *P*(*H*10) = *P*(*S*1)·β.

Вероятность пропуска дефекта равна вероятности произведения двух событий: наличие неисправного состояния и значения . Тогда

3.5. Определяется средний риск *R*

.

4. Результаты решения задачи представить в виде таблицы 3:

Таблица 3. Определение граничного значения температуры *х*гр и среднего риска *R*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | Метод | Граничное значение | Вероятность ложной  тревоги | Вероятность пропуска  дефекта | Средний риск |
| 1 | Метод минимального среднего риска |  |  |  |  |
| 2 | Метод минимальной вероятности ошибочного решения |  |  |  |  |

5. Сделать выводы по работе.

6. Оформить и защитить отчет.