Міністерство освіти і науки України
Донецький національний технічний університет

Кафедра "Вища математика"

Збірник науково-методичних робіт

Випуск 2

Донецьк -2004

УДК 512.643, 517. 944(09), 517.926, 519.61/.64, 531.38, 535.36, 539.238, 622.831.

Рекомендовано до друку Радою Донецького Національного технічного Університету
Протокол № 4 від 30. 05. 2003 р.

Збірник науково-методичних робіт. - Вип. 2. - Донецьк: ДонНТУ, 2004. - 200с.

Процеси гуманізації й гуманітаризації освітньої системи в Україні передбачають виконання значної кількісті суттєвих вимог щодо організації навчального процесу у вищих навчальних закладах. Відповідно до цього виникає нагальна потреба в особистісній зорієнтованості навчання, а саме - в створенні умов для розвитку позитивних, і в першу чергу потенцій кожного студента.

В збірнику представлено результати науково-методичних досліджень, в яких обгрунтовуються нові підходи до певних питань методики викладання вищої математики, досліджено окремі історичні аспекти розвитку матетематики, розглянуто низку цікавих задач з затосування математики в різних галузях науки й техніки.

Адреса редакційної колегії : Україна, 83050, м. Донецьк, вул. Артема, 96, ДонНТУ, 3-й учбовий корпус, кафедра "Вища математика", тел. (0622) 999901.

УДК 622.232

О выборе уровня значимости при проверке статистических гипотез

С. Г. Ехилевский, В. В. Малашенко
Донецкий национальный технический университет

В роботі показано, що істинною незалежною змінною, що задає об'єм вибірки та критичне значення критерію, є ризик. Для обґрунтованого вибору рівня значущості необхідно конкретизувати закон розподілу критерію.

Как известно, статистической называют гипотезу о виде закона распределения случайной величины и его параметрах. Принять или отвергнуть гипотезу нужно на основе данных выборки, что вносит элемент случайности и может привести к ошибкам (принятию неверной гипотезы или неприятию верной). Поэтому можно говорить лишь о той или иной вероятности того, что гипотеза имеет (или не имеет) место.

Для принятия решения из данных выборки составляют некоторую случайную величину (критерий), возможные значения которой в зависимости от объема выборки и требуемой вероятности делят на две части. Внутри одной из них гипотезу принимают, а во второй –отвергают.

Пример. Пусть гипотеза состоит в том, что матожидание нормально распределенной случайной величины равно числу m. Критерием в этом случае является выборочное среднее, а областью принятия – некоторая δ-окрестность m, найденная с помощью имеющихся таблиц и отвечающая доверительной вероятности γ.

Следует подчеркнуть, что при этом, во-первых, остается конечная вероятность $\alpha = 1 - \gamma$ отклонить правильную гипотезу, называемая уровнем значимости. Во-вторых, само принятое решение может измениться при увеличении объема выборки (δ-окрестность уменьшится, и критерий перестанет в нее попадать).

Таким образом, сразу возникает два вопроса:

1.Из каких соображений выбирать α (а значит и границы окрестности, являющиеся критическими значениями критерия)?

2.Каким объемом следует ограничить выборку?

С ответами на них непосредственно связан метод минимума риска, излагаемый, однако, в литературе вне данного контекста [1].

Метод минимума риска.

Пусть ущерб, связанный с ошибкой первого рода (отклонение правильной гипотезы) равен C_1, а второго рода (принятие неправильной) – C_2 [1]. Тогда матожидание ущерба C, называемое риском, равно

$$r = M(C) = P_1 \cdot C_1 + P_2 \cdot C_2, \qquad (1)$$

где

$$P_1 = P(H) \cdot P_H(\bar{A}) \qquad (2)$$

- вероятность совершить ошибку первого рода, в которой $P(H)$ – вероятность того, что проверяемая гипотеза имеет место, а $P_H(\overline{A})$ – вероятность при этом условии отклонить гипотезу (именно $P_H(\overline{A})$ при данном объеме выборки связана с α). Аналогично вероятность совершить ошибку второго рода представим в виде[2]

$$P_2 = P(H) \cdot P_H(A) \qquad (3)$$

Естественно при данном объеме выборки выбирать α так, чтобы риск был минимален. Если даже в этом случае он оказывается неприемлемо велик, следует увеличивать объем выборки, принимая однако во внимание растущие затраты на сбор и обработку статистических данных.

Задача. Электролампы поступают партиями с двух заводов. Вероятность выпуска бракованной продукции первым и вторым заводом равна соответственно p_1 и p_2. Известно, что 1/5 всех партий поступает с первого завода. Требуется на основании данных выборочного контроля определить, на каком заводе изготовлена выставленная на продажу партия.

Решение. Пусть n - объем выборки. Значит, число бракованных ламп y – случайная величина с возможными значениями 0,1,2,…,n. Очевидно, что y в данном случае является критерием согласия. Положим для определенности $p_1 > p_2$. Тогда, если $y > y^*$, считаем, что партия изготовлена на первом заводе. Здесь y^* - критическое значение критерия, которое предстоит определить.

Нетривиальность задачи состоит в том, что из плохой партии можно случайно осуществить выборку качественных изделий, и даже первый завод может случайно выпустить хорошую партию. Поэтому партию первого

---

[1] Определение C_1 и C_2 – задача специалистов той отрасли, которая прибегает к услугам матстатистики.

[2] Заметим, что гипотез может быть много, однако все они, кроме H, могут быть объединены в альтернативную.

завода можно счесть изготовленной на втором и наоборот. Коротко эти ошибки будем называть первой и второй соответственно. Найдем их вероятности. По схеме Бернулли имеем следующие законы распределения *y* для первого и второго заводов

$$P_n^{(i)}(y) = C_n^y \cdot p_i^y \cdot (1 - p_i)^{n-y}, \quad (i = 1, 2) \qquad (4)$$

где C_n^y - число сочетаний из n по *y*. Пусть гипотеза H_1 заключается в том, что партия изготовлена на первом заводе. В соответствии с ранее изложенным вероятность отклонить H_1 при условии, что она верна, равна сумме[3]

$$P_{H_1}(A_1) = \sum_{y \le y^*} P_n^{(1)}(y). \qquad (5)$$

Случайное событие A_i состоит в том, что допущена i-я ошибка. Аналогично для второй гипотезы имеем

$$P_{H_2}(A_2) = \sum_{y > y^*} P_n^{(2)}(y) = 1 - \sum_{y \le y^*} P_n^{(2)}(y).$$

(6)

Здесь учтено, что полная вероятность равна единице. По условию априорные вероятности гипотез задаются равенствами

$$P(H_1) = 1/5, \; P(H_2) = 4/5. \qquad (7)$$

И, наконец, по теореме умножения получаем вероятности ошибок

$$P(A_i) = \sum_{y \le y^*} P(H_i) \cdot P_{H_i}(A_i). \qquad (8)$$

Таким образом, для риска (см. (1)) имеем

$$r = C_1 \cdot P(H_1) \cdot \sum_{y \le y^*} P_n^{(1)}(y) + C_2 \cdot P(H_2) \cdot (1 - \sum_{y \le y^*} P_n^{(2)}(y)) =$$

$$= C_2 \cdot P(H_2) + \sum_{y \le y^*} (C_1 \cdot P(H_1) \cdot P_n^{(1)}(y) - C_2 \cdot P(H_2) \cdot P_n^{(2)}(y)).$$

(9)

[3] В(5) складываются вероятности хороших выборок (y < y*) из плохой партии (i =1), найденные в рамках схемы Бернулли.

Чтобы при данном n он был минимален, суммировать в (9) следует до тех пор, пока выражение в квадратных скобках

$$f(y) = C_1 \cdot P(H_1) \cdot P_n^{(1)}(y) \ - \ C_2 \cdot P(H_2) \cdot P_n^{(2)}(y).$$

(10)

- отрицательно[4]. Из условия f(y)=0 и определим критическое значение критерия

$$y^* = \frac{\ln\left[\dfrac{C_1 \cdot P(H_1) \cdot (1-p_1)^n}{C_2 \cdot P(H_2) \cdot (1-p_2)^n}\right]}{\ln\left[\dfrac{p_2 \cdot (1-p_1)}{p_1 \cdot (1-p_2)}\right]} \equiv y^n(n). \qquad (11)$$

Выясним, каким должен быть объем выборки. Для малых n может оказаться, что y^* отрицательно или больше n, если очень велик ущерб, связанный соответственно с первой или второй ошибками. В этом легко убедиться, рассмотрев в (11) предельные ситуации с C_1/C_2 , стремящимся к бесконечности или к нулю. В первом случае, чтобы меньше рисковать, партию нужно считать худшей (принимать H_1),даже если оказалось, что y = 0. А во втором – качественной даже при y = n. Очевидно, что делать такие выборки бессмысленно. В связи с этим определим минимальное разумное n. Заметим для этого, что логарифм в знаменателе (11) отрицателен. Поэтому первая ситуация (y* < 0) не возникнет, если

$$\frac{C_1 \cdot P(H_1) \cdot (1-p_1)^n}{C_2 \cdot P(H_2) \cdot (1-p_2)^n} \leq 1. \qquad (12)$$

А чтобы не возникла вторая ситуация, (y* > n), потребуем выполнения неравенства

---

[4] Если n достаточно велико, то f(0) < 0, а f(n) > 0, в чем легко убедиться с помощью (4),(10):

$$f(0) = C_1 \cdot P(H_1) \cdot (1-p_1)^n - C_2 \cdot P(H_2) \cdot (1-p_2)^n \xrightarrow[n\to\infty]{} - C_2 \cdot P(H_2)$$

т.к. $p_1 > p_2$. По той же причине

$$f(n) = C_1 \cdot P(H_1) \cdot p_1^{\,n} - C_2 \cdot P(H_2) \cdot p_2^{\,n} \xrightarrow[n\to\infty]{} C_1 \cdot P(H_1) \cdot p_1^{\,n}.$$

$$\frac{\ln\left[\dfrac{C_1\cdot P(H_1)\cdot(1-p_1)^n}{C_2\cdot P(H_2)\cdot(1-p_2)^n}\right]^n}{\ln\left[\dfrac{p_2\cdot(1-p_1)}{p_1\cdot(1-p_2)}\right]}\le n. \qquad (13)$$

В зависимости от конкретных C_1, C_2, $P(H_1)$ и $P(H_2)$ пользоваться нужно более сильным из условий (12), (13).

Из формул (9), (11) следует, что в случае равенства (12) $r = C_2\cdot P(H_2)$, а в случае равенства (13) $r = C_1\cdot P(H_1)$.[5] Для больших n при оптимальном выборе y* получим соотношение

$$r = C_2\cdot P(H_2)+\sum_{y=0}^{y^*(n)} f(y)\equiv r(n), \qquad (14)$$

из которого, задавшись приемлемым значением риска, можно найти достаточный для этого объем выборки. Естественно, он будет тем меньше, чем больше отличие между p_1 и p_2. В предельной ситуации, когда первый завод выпускает сплошной брак, а второй – только качественную продукцию ($p_1 - p_2 = 1$), происхождение партии достоверно определяется контролем единственного изделия (r = 0 уже для n=1). Последнее совершенно очевидно, однако с целью проверки развитого формализма убедимся в этом аналитически. Подставив в (11) p_1 и p_2, равные соответственно $1-\varepsilon$ и ε , в пределе, когда ε стремится к нулю, получим, что y* = n/2.Таким образом для n = 1 целочисленное по смыслу y может принять в сумме (9) только нулевое значение. В результате с учетом (4) получим

$$r = C_2\cdot P(H_2)+C_1\cdot P(H_1)\cdot P_1^{(1)}(0)-C_2\cdot P(H_2)\cdot P_1^{(2)}(0)=$$

$$= C_2\cdot P(H_2)+C_1\cdot P(H_1)\cdot\varepsilon - C_2\cdot P(H_2)\cdot(1-\varepsilon)\underset{\varepsilon\to 0}{\longrightarrow} 0,$$

что и требовалось доказать.

С помощью (11) нетрудно получить асимптотическую формулу

[5] В последнем нетрудно убедиться, заметив в (9), что

$$\sum_{y=0}^{n} P_n^{(i)}(y)\equiv 1..$$

$$y^*(n) \xrightarrow[n\to\infty]{} n \frac{\ln\left(\dfrac{1-p_1}{1-p_2}\right)}{\ln\left(\dfrac{p_2(1-p_1)}{p_1(1-p_2)}\right)} < n\,. \qquad (15)$$

Т.е., с ростом n число отрицательных слагаемых в (9) увеличивается, но сами они уменьшаются по модулю, ибо r > 0 по смыслу. Тем не менее при больших n риск сводится к нулю,в чем, однако, очень трудно убедиться аналитически из-за необходимости частичного суммирования в (14) биномиальных разложений с очень большим числом членов.

Обсудим теперь проблему выбора уровня значимости. Его определение, приведенное выше, не совсем удачно с методической точки зрения, ибо правильной по воле случая может оказаться любая из альтернативных гипотез. Чтобы избежать двусмысленности, уровнем значимости данной гипотезы назовем условную вероятность ее отвергнуть при условии, что она (данная гипотеза) верна. В рамках такого определения уровни значимости первой и второй гипотез вычисляются по формулам (5), (6)[6]. Впрочем, сделать это можно лишь из чистого любопытства, т. к. истиной независимой переменной, задающей объем выборки и критическое значение критерия, является приемлемый риск.

Укажем в этой связи, что учебные задачи, в которых в той или иной форме предлагается с помощью готовых таблиц проверить, укладываются ли данные n, аль и некоторое значение критерия в рамки определенной гипотезы, являются некорректными, поскольку нет соображений, позволяющих выбрать уровень значимости, если для альтернативной гипотезы не конкретизирован закон распределения критерия.

Литература

[6] Заметим, что, несмотря на альтернативность гипотез, сумма их уровней значимости не равна единице

$$\alpha_1 + \alpha_2 = P_{H1}(A_1) + P_{H2}(A_2) = 1 - \sum_{y<y^*}\left(P_n^{(1)}(y) - P_n^{(2)}(y)\right).$$

Объяснение заключается в том, что разным гипотезам соответствуют разные законы распределения критерия. В нашем случае $P_n^{(1)}(y) \neq P_n^{(2)}(y)$, т. к. $p_1 \neq p_2$.

1. Гнеденко Б.В. Курс теории вероятностей.-М.:Наука, 1969.-400 с.

СОДЕРЖАНИЕ