

4.8. Треугольный эксперимент. Более простая процедура, применяемая на практике (часто во время коммерческих презентаций), называется *треугольным экспериментом*.

Предположим, что три стакана наполнены одинаковыми по цвету напитками. В два из них налит напиток D_1 , а в третий — напиток D_2 . Выбирая стакан наугад, экспериментатор предлагает желающим попробовать содержимое и определить его содержимое. Если ответы даются случайным образом, то при большой выборке только половина респондентов угадает D_1 . Если же D_1 имеет отличительные вкусовые особенности, то примерно две трети ответов будут правильными.

4.9. Последовательный анализ. Планирование эксперимента представляет собой в настоящее время обширную область теоретических и практических результатов. Среди многообразия его методов рассмотрим еще только один пример из той его части, которая называется *последовательным анализом*.

Предположим, что необходимо оценить некоторую величину θ , которую можно измерить с ошибкой u , то есть при каждом измерении мы получаем значения случайной величины $X = \theta + u$, где u напрямую не наблюдается. При этом известно, что $E[u] = 0$ и $\text{var}[u] = \sigma^2$. Если проведены n измерений (их результаты обозначим X_1, \dots, X_n), то можно вычислить выборочное среднее

$$(6) \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \theta + \frac{1}{n} \sum_{i=1}^n u_i.$$

Именно выборочное среднее \bar{X} мы и выбираем в качестве оценки для θ . Каким следует выбрать объем выборки n ? В традиционной статистике этот вопрос заканчивают словами

“... чтобы оценка была как можно лучше”. Поскольку

$$E[\bar{X}] = \theta, \quad \text{var}[\bar{X}] = \frac{\sigma^2}{n},$$

то с ростом n точность оценки (6) становится выше и поэтому ответ на заданный вопрос очевиден: n надо выбрать как можно бóльшим. Этот подход игнорирует затраты на проведение каждого эксперимента и становится неприемлемым, если их приходится учитывать.

Рассмотрим другой подход для определения оптимального n , который учитывает потери от проведения эксперимента, а также потери от использования выбранной оценки. Пусть $L(\theta, t)$ — это цена, которую приходится платить за использование значения t вместо настоящего значения θ и пусть $C(n)$ — это цена за использование выборки размера n . В частном случае, когда

$$(7) \quad L(\theta, t) = k(t - \theta)^2,$$

$$(8) \quad C(n) = cn$$

для некоторых коэффициентов k и c , *ожидаемые потери* R (*ожидаемая цена*) за использование оценки \bar{X} вместо θ для выборки размера n равны

$$(9) \quad R(\theta, n) = kE[(\bar{X} - \theta)^2] + cn = \frac{k\sigma^2}{n} + cn.$$

Минимизируя правую часть по n (как будто бы n не дискретный, а непрерывный параметр), находим точку минимума

$$(10) \quad n_0 = \sqrt{k\sigma^2/c},$$

при которой потери равны $R(\theta, n_0) = 2(ck\sigma^2)^{1/2}$.

Замечание 2. Задача поиска оптимального n для функции (9) решена приблизительно с помощью дифференцирования по параметру n , хотя на самом деле он дискретный. Так как этот метод довольно прост, он часто используется для нахождения приближения к настоящему решению, тем более, что это приближение обычно приемлемо. Например, пусть $k = \sigma^2 = c = 1$. Тогда “оптимальное” значение n_0 , найденное в (10), равно 1. С другой стороны, мы можем решить задачу минимизации функции (9) точно. Действительно,

$$R(\theta, n) = n + \frac{1}{n} \geq 2, \quad \text{так как} \quad n + \frac{1}{n} - 2 = \left(\sqrt{n} - \frac{1}{\sqrt{n}} \right)^2.$$

Таким образом, минимум функции $R(\theta, n)$ действительно достигается при $n = 1$.

Впрочем, этот вывод можно было бы сделать и без вычислений, так как “оптимальное” значение, найденное согласно (10), является натуральным числом, поскольку

$$\min_t R(\theta, t) \leq \min_n R(\theta, n) \leq R(\theta, n_0)$$

для любого натурального числа n_0 .

Эта задача порождает несколько важных вопросов.

1. *Почему выбраны именно такие функции цены?* Функция (8), точнее $c_0 + cn$ при некотором c_0 , считается общепринятой при проведении экспериментов: она отражает тот факт, что цена каждого эксперимента такая же, как и цена любого другого. Параметр c_0 не сказывается на последующих вычислениях оптимального значения n .

Квадратичная *функция потерь* (7) является классической. Она считается приемлемой аппроксимацией для многих других (более общих) функций $L(\theta, t)$, которые равны

0 при $t = \theta$ и имеют минимум в этой точке. Для таких функций

$$L(\theta, t)|_{t=\theta} = 0, \quad L'_t(\theta, t)|_{t=\theta} = 0.$$

Раскладывая в ряд Тейлора по второму параметру, имеем в малой окрестности точки $t = \theta$,

$$\begin{aligned} L(\theta, t) &= L(\theta, t)|_{t=\theta} + L'_t(\theta, t)|_{t=\theta}(t - \theta) \\ &\quad + \frac{1}{2}L''_t(\theta, t)|_{t=\theta}(t - \theta)^2 + \dots \\ &\approx k(t - \theta)^2. \end{aligned}$$

Коэффициент k в этой аппроксимации зависит от θ , но мы считаем, что он не сильно меняется в выбранной окрестности.

2. Почему мы выбрали средний риск при выборе n ? Этот вопрос действительно важен. Если бы мы решали задачу несколько раз, то выбор среднего в качестве характеристики для оптимизации был бы обоснован. Но ведь мы решаем задачу только один раз!

Выбор среднего значения можно объяснить в терминах *функций полезности*. При очень слабых предположениях, предпочтения индивидуума выражаются через его функцию полезности, которая измеряет полезность каждого варианта для выбранного индивидуума. Например, в лотерее имеются два возможных варианта (“успех” и “неудача”), полезность которых u_1 и u_2 , а вероятности p и $1 - p$. Полезностью лотереи для данного индивидуума является

$$u = pu_1 + (1 - p)u_2.$$

Иными словами, полезность случайного варианта равна математическому ожиданию этого варианта. Поэтому, если

измерять потери в терминах полезности, то математическое ожидание $L(\theta, \bar{X})$ является подходящей характеристикой для потерь, связанных с проведением n экспериментов.

3. Почему мы выбрали \bar{X} в качестве оценки? Выбор оценки для θ зависит от наших знаний о распределении u . Если распределение нормально, то \bar{X} очень хорошая оценка. Однако в некоторых других случаях существуют и лучшие, чем \bar{X} , оценки.

4. Что делать, если σ^2 неизвестно? Одно из решений состоит в замене σ^2 ее оценкой. Поскольку $R(\theta, n)$ не сильно меняется при малых изменениях n в окрестности n_0 , то такая подмена не должна сильно сказываться на потерях. Можно действовать последовательно, а именно: получив наблюдение n , оцениваем σ^2 выборочной дисперсией

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Если

$$n \geq \sqrt{ks^2/c},$$

то заканчиваем эксперименты.

Таким образом, мы заранее не знаем, когда надо заканчивать эксперименты. Решение об окончании мы принимаем, последовательно проводя эксперименты один за другим. Именно это и оюясняет термин “последовательный анализ”.

У П Р А Ж Н Е Н И Я

Упражнение 1. Доказать, что (3) — это латинский квадрат.

Упражнение 2. Доказать теорему 1.