Основные определения теории проверки гипотез

Пусть имеется выборка $X_1, X_2, \ldots X_n$ из генеральной совокупности с неизвестной теоретической функцией распределения.

Определение: *Статистической гипотезой* называется любое предположение о виде теоретической функции распределения, то есть статистическая гипотеза – это рассматриваемое предположение о величине параметра генеральной совокупности.. Имеются две непересекающиеся гипотезы: $\boldsymbol{H_0}$ и $\boldsymbol{H_1}$. $\boldsymbol{H_0}$ – нулевая (основная) гипотеза, $\boldsymbol{H_1}$ – альтернативная (конкурирующая) гипотеза. Принято считать, что $\boldsymbol{H_0}$ –гипотеза о сходстве, H_1 –гипотеза о различии.

- **Нулевая гипотеза** – это допущение, которое считается верным до тех пор, пока не будет доказано обратное, исходя из результатов статистической проверки.
- **Альтернативная гипотеза** – это гипотеза, которая принимается, если в результате статистической проверки отвергается нулевая гипотеза.

Основные определения теории проверки гипотез

Определение: *Статистическим критерием (тестом)* называется правило, позволяющее на основании наблюдений $X_1, X_2, \ldots X_n$ принять нулевую гипотезу H_0 или отвергнуть ее в пользу альтернативной H_1.

Проверка гипотезы может быть односторонней или двусторонней.

Определение: *Односторонний критерий* используется в тех случаях, когда необходимо знать, является ли параметр генеральной совокупности > (правосторонний критерий) или < (левосторонний критерий) предполагаемого значения.

Определение: *Двусторонний критерий* используется в тех случаях, когда интересует, отличаются ли реальные значения параметра от предполагаемого значения.

Определение: *Критическую область* составляют те значения выборочных статистических показателей, которые ведут к отказу от нулевой гипотезы.

Уровень значимости

Определение: *Уровень значимости* – вероятность ошибочного отклонения нулевой гипотезы H_0 (вероятность ошибки I рода). При статистическом анализе исследователь должен выбрать необходимый уровень значимости. При этом считают низшим уровнем значимости значение α=0.05, достаточным уровнем - α=0.01, высшем уровнем α =0.001.

Иногда, *доверительной вероятностью* считается величина $p=1-\alpha$

Возможные решения статистического критерия:

Результат проверки гипотезы	***Возможные состояния проверяемой гипотезы***	
	Верна гипотеза H_0	Верна гипотеза H_1
H_0 отклоняется	Ошибка I рода	Правильное решение
H_0 не отклоняется	Правильное решение	Ошибка II рода

Ошибки I и II рода

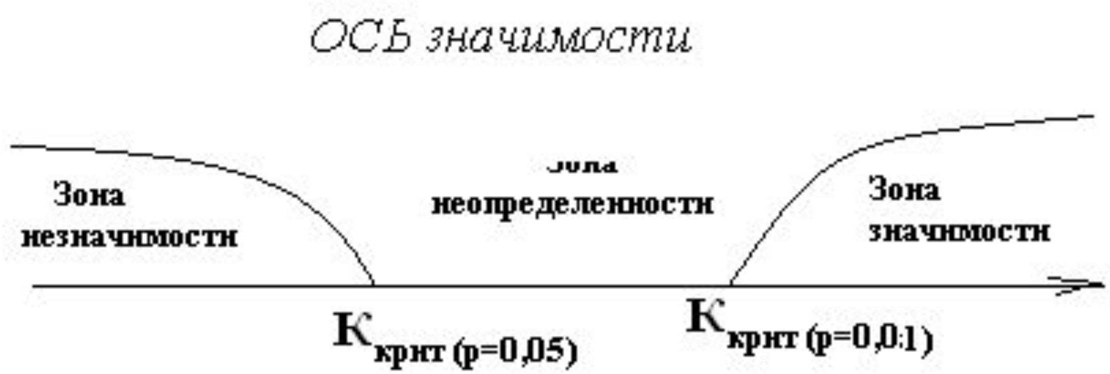
Определение 1: В процессе проверки гипотезы существует вероятность того, что H_0 будет отвергнута, когда в действительности она должна быть принята. Это называется ***ошибкой первого рода.*** Вероятность допущения ошибки первого рода это уровень значимости. Таким образом, когда выбирают 5% уровень значимости для проверки, одновременно допускают, что в 5% случаев должны отвергнуть H_0, хотя она и верна.

Определение 2: Второй вид ошибок имеет место при принятии нулевой гипотезы, в то время как в действительности она должна быть отвергнута. такая ошибка называется ***ошибкой второго рода***.

		Действительность	
		H_0 *- верна* Обвиняемый невиновен	H_0 *- ложна* Обвиняемый виновен
РЕШЕНИЕ	H_0 *- принята* Обвиняемый освобожден	X	Ошибка II рода
	H_0 *- отвергнута* Обвиняемый наказан	Ошибка I рода	X

Этапы принятия статистического решения

1. Формулировка нулевой и альтернативной гипотез.
2. Определение объема выборки.
3. Выбор соответствующего уровня значимости или вероятности отклонения гипотезы H_0 ($p \leq 0.05$).
4. Выбор статистического метода, который зависит от типа решаемой задачи.
5. Вычисление значения *выборочной статистики* $K_n = K_n(X_1, X_2, \ldots X_n)$ на основании наблюдений $X_1, X_2, \ldots X_n$.
6. Если гипотеза H_0 верна, то распределение случайной величины K_n известно (затабулировано). Нахождение по таблице для выбранного статистического метода *критической области* K_α для определенного уровня значимости.
7. Сравнение эмпирического и критического значений. Если $K_n \in K_\alpha$, то принимается H_0; если $K_n \notin K_\alpha$, то H_0 отвергается в пользу альтернативной.
8. Формулировка принятия решения (выбор гипотезы H_0 или H_1).



При попадании выборочной статистики в зону незначимости принимается гипотеза H_0 об отсутствии различий. В случае попадания в зону значимости принимается гипотеза H_1 о наличии различий, а гипотеза H_0 отклоняется. При попадании выборочной статистики в зону неопределенности в зависимости от важности решаемой задачи можно принять H_1 на уровне 5% или принять H_0 на 1% уровне. В этом случае можно допустить ошибки I или II рода. В этих обстоятельствах лучше увеличить объем выборки.

Проверка гипотезы о соответствии исправленной выборочной дисперсии величине генеральной дисперсии нормальной совокупности

Стандартизированный статистический критерий (тест) для проверки такой гипотезы рассчитывается как:

$$\chi^2_{расч} = \frac{(n-1)\cdot S^2}{\sigma_0^2}, \qquad \textbf{(1)}$$

где σ_0^2 – проверяемое значение генеральной дисперсии,
а S^2 – исправленная выборочная дисперсия, n – объем выборки.

Левосторонняя проверка: нулевая и альтернативная гипотезы имеют вид:

H_0: $S^2=\sigma^2$ – равенство неизвестной генеральной дисперсии S^2;
H_1: $S^2<\sigma^2$.

Правило принятия решения: принять H_0, если $\chi^2_{расч} > \chi^2_{1-\alpha,k}$,
отвергнуть H_0, если $\chi^2_{расч} > \chi^2_{\alpha,k}$

Здесь α – уровень значимости принятия гипотезы, $k=n-1$ – число степеней свободы $\chi^2_{1-\alpha,k}$ - определяется по таблице χ^2–распределения.

Проверка гипотезы о соответствии исправленной выборочной дисперсии величине генеральной дисперсии нормальной совокупности

Правосторонняя проверка: нулевая и альтернативная гипотезы имеют вид:

H_0: $S^2=\sigma^2$ – равенство неизвестной генеральной дисперсии S^2;

H_1: $S^2>\sigma^2$.

Правило принятия решения: принять H_0, если $\chi^2_{расч} < \chi^2_{\alpha,k}$,

отвергнуть H_0, если $\chi^2_{расч} > \chi^2_{\alpha,k}$.

Двусторонняя проверка: нулевая и альтернативная гипотезы имеют вид:

H_0: $S^2=\sigma^2$ – равенство неизвестной генеральной дисперсии S^2;

H_1: $S^2\neq\sigma^2$.

Правило принятия решения: принять H_0, если $\chi^2_{1-\frac{\alpha}{2},k} < \chi^2_{расч} < \chi^2_{\frac{\alpha}{2},k}$,

отвергнуть H_0 в противном случае.

Проверка гипотезы о соответствии выборочной средней величине генеральной средней нормальной совокупности

Формируем гипотезы о равенстве генеральной $\boldsymbol{\mu}$ и выборочной средней $\boldsymbol{\mu_0}$.

H_0: $\mu=\mu_0$;

H_1: $\mu\neq\mu_0$.

<u>**Правило**</u> **принятия решения**: принять H_0, если $-z_{крит} < \dfrac{\bar{x}-\mu_0}{\sqrt{S^2/n}} < z_{крит}$,

в противном случае принять H_1.

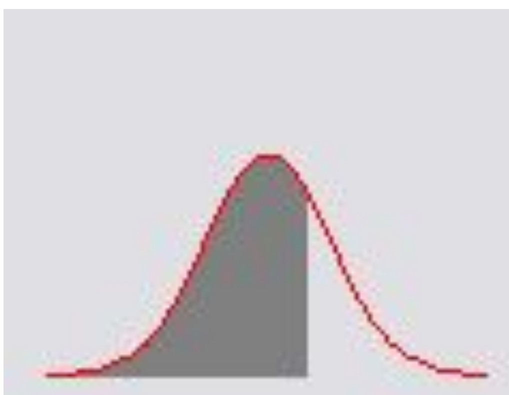
$Z_{крит}$ определяется из таблиц функции Лапласа

из равенства $\boldsymbol{\Phi(z_{крит})=(1-\alpha)/2}$.

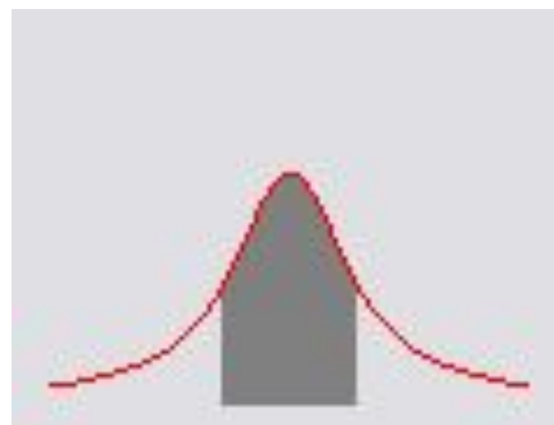
Метод p-value

Величина *p* – это значение, которое в случае верности нулевой гипотезы представляет собой вероятность получения величины стандартизированного критерия проверки, большего по абсолютному значению, чем рассчитанный критерий проверки.

В случае односторонней проверки Р равно площади под кривой слева (левосторонняя проверка) или справа 9правосторонняя проверка) от значения критерия проверки. В случае двусторонней проверки она равна удвоенной площади в части под кривой справа или слева от критерия проверки.



Односторонняя проверка



Двусторонняя проверка

Метод p-value

В методе p-value правило принятия решения одинаково независимо то того, выполняется левосторонняя, правосторонняя или двусторонняя проверка. Обозначив степень значимости для проверки через α, получим следующее **правило принятия решения**:

- Принять H_0, если p-value $\geq \alpha$
- В противном случае, отвергнуть H_0.

Расчет величины p:

Для того чтобы найти величину p, прежде всего рассчитывают стандартный критерий проверки, а затем, зная число степеней свободы, находят вероятности (площади в граничных областях), соответствующие показателям статистики (F или t или z), которые охватывают снизу и сверху рассчитанный критерий проверки. После этого с помощью интерполяции, исходя из полученных вероятностей, находят величину p.

Задача оценивания

Пусть имеются данные выборки, например значения некоторого признака, $X_1, X_2, \ldots, X_n$, полученные в результате n наблюдений. Для того чтобы найти **статистическую оценку** θ неизвестного параметра теоретического распределения через эти данные необходимо найти функцию от наблюдаемых случайных величин, которые дают приближенное значение оцениваемого параметра.

Статистическую оценку, которая определяется одним числом, называют **точечной**.

Свойства оценок

Полученные оценки должны быть **достоверными**, т.е. обладать свойствами ***несмещенности, эффективности и состоятельности***.

- **Несмешанной** называют статистическую оценку θ^*, математическое ожидание которой равно оценивающему параметру θ при любом объеме выборки, т.е. $M(\theta^*)=\theta$.
- **Эффективной** оценкой называют статистическую оценку θ^*, которая при заданном объеме выборки n имеет наименьшую возможную дисперсию.
- **Состоятельной** называют статистическую оценку, которая при $n\to\infty$ и стремится по вероятности к оцениваемому параметру, т.е.:

.

$$\lim_{n\to\infty} P(\theta^*)=\theta$$

Метод моментов для точечной оценки параметра распределения

Оценка одного параметра

Вид плотности распределения $f(x, \theta)$.

Требуется найти точечную оценку $\hat{\theta}$.

Для оценки *одного параметра* ***достаточно*** *одного уравнения*, относительного этого параметра.

Пусть $M(x) = \bar{x}_в$

Тогда $M(x) = \int\limits_{-\infty}^{\infty} x \cdot f(x,\theta)dx = \varphi(\theta)$ $\Rightarrow$ $\varphi(\theta) = \bar{x}_в$

Решив уравнение относительно параметра θ , найдем точечную оценку $\hat{\theta}$

Следовательно оценка есть функция от вариант выборки:

$$\hat{\theta} = \psi(x_1, x_2, ..., x_n)$$

Метод моментов для точечной оценки параметра распределения

Оценка двух параметров

Вид плотности распределения $f(x, \theta_1, \theta_2)$.

Требуется найти точечные оценки $\hat{\theta}_1$ и $\hat{\theta}_2$

Для оценки *двух параметров* ***достаточно*** *системы двух уравнений*, относительного этих параметров.

Пусть $\begin{cases} M(x) = \bar{x}_в \\ D(x) = D_в \end{cases}$

Тогда $$\begin{cases} M(x) = \int\limits_{-\infty}^{\infty} x \cdot f(x, \theta_1, \theta_2) dx = \varphi_1(\theta_1) \\ D(x) = M(x^2) - [M(x)]^2 = \varphi_2(\theta_2) \end{cases}$$

Решив систему относительно параметров θ_1, θ_2 , найдем точечные оценки $\hat{\theta}_1$ $\hat{\theta}_2$

Следовательно оценки есть функции от вариант выборки:

$$\hat{\theta}_1 = \psi_1(x_1, x_2, ..., x_n)$$
$$\hat{\theta}_2 = \psi_2(x_1, x_2, ..., x_n)$$

Метод максимального правдоподобия

Для дискретных случайных величин.

Пусть X дискретная случайная величина, которая принимает возможные значения x_1, $x_2,\ldots,x_n$. Пусть закон распределения задан, но неизвестен параметр распределения θ. Требуется найти точечную оценку $\hat{\theta}$.

Вероятность того, что величина X, примет значение x_i, $p(x_i, \theta)$.

Определение: Функцией правдоподобия дискретной случайной величины X называют функцию аргумента θ

$$L(x_1, x_2, ..., x_n; \theta) = p(x_1;\theta)\,p(x_2;\theta)\cdot ... \cdot p(x_n;\theta) = \prod_{i=1}^{n} p(x_i;\theta)$$

Где x_1, $x_2,\ldots,x_n$. – фиксированные числа.

Определение: Логарифмической функцией правдоподобия дискретной случайной величины X называют функцию аргумента θ

$$l = \ln L(x_1, x_2, ..., x_n; \theta) = \sum_{i=1}^{n} p(x_i;\theta)$$

Метод максимального правдоподобия

Определение: Оценкой максимального правдоподобия называют такую оценку $\hat{\theta}$, для которой функция правдоподобия достигает максимума.

$$\hat{\theta} = \max_{i=1} L(x_1, x_2, ..., x_n; \theta) = \max_{i=1} l(x_1, x_2, ..., x_n; \theta)$$

Для ее нахождения решают уравнение, называемое уравнением правдоподобия: $\dfrac{d \ln L}{d\theta} = 0$

Если при $\theta = \hat{\theta}$, $\dfrac{d^2 \ln L}{d\theta^2} < 0$, то $\hat{\theta}$ - точка максимума.

Метод максимального правдоподобия

Для непрерывных случайных величин.

Пусть X непрерывная случайная величина, которая пв результате испытания приняла значения x_1, $x_2, \ldots, x_n$. Пусть вид плотности распределения $f(x)$ известен, но неизвестен параметр распределения θ. Требуется найти точечную оценку $\hat{\theta}$.

Определение: Функцией правдоподобия непрерывной случайной величины X называют функцию аргумента θ:

$$L(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \cdot \ldots \cdot f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

Где $x_1, x_2, \ldots, x_n$. – фиксированные числа.

Оценку максимального правдоподобия неизвестного параметра распределения непрерывной случайной величины ищут также, как и в случае с дискретной случайной величины.

Метод максимального правдоподобия

Для непрерывных случайных величин.

Если плотность распределения $f(x)$ непрерывной случайной величины X определяется двумя неизвестными параметрами θ_1, θ_2, то функция правдоподобия является функцией двух аргументов θ_1, θ_2:

$$L(x_1, x_2, ..., x_n; \theta_1, \theta_2) = f(x_1; \theta_1, \theta_2) f(x_2; \theta_1, \theta_2) \cdot ... \cdot f(x_n; \theta_1, \theta_2) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2)$$

Где $x_1, x_2, \ldots, x_n$. – фиксированные числа.

Для нахождения параметров θ_1, θ_2 решают систему уравнений:

$$\begin{cases} \dfrac{\partial \ln L}{\partial \theta_1} = 0, \\ \dfrac{\partial \ln L}{\partial \theta_2} = 0. \end{cases}$$

Статистическая задача оценивания

Задача: по наблюдениям x_1, $x_2,\dots,x_n$ над случайной величиной X, распределенной равномерно на отрезке [0, a], оценить неизвестный параметр a.

Сравним три способа оценивания:

1.Метод моментов $\hat{a}_1 = \frac{2}{n}\sum_{i=1}^{n} x_i$

2.Метод максимального правдоподобия $\hat{a}_2 = \frac{n+1}{n}\max_i x_i$

3.Метод порядковых статистик $\hat{a}_3 = 2\hat{x}_{0,5} = x_{(k)} + x_{(k+1)}$

Где $\hat{x}_{0,5} = \frac{x_{(k)} + x_{(k+1)}}{2}$ - выборочная квантиль порядка 0,5, т.е. выборочная медиана, $x_{(k)}$ - член вариационного ряда с номером k. (причем n=2k).

Теоретическое сравнение оценок

1. Все оценки $\hat{a}_1, \hat{a}_2, \hat{a}_3$ являются несмещенными, их математические ожидания равны истинным параметрам a. (***доказать сам-но***)
2. Дисперсии оценок: (***будет доказано на лекции***)

$$D\hat{a}_1 = D\left(\frac{2}{n}\sum_{i=1}^{n} x_i\right) = \frac{a^2}{n} \qquad D\hat{a}_2 = D\left(\frac{n+1}{n}\max_i x_i\right) = \frac{a^2}{n(n+2)}$$

$$D\hat{a}_3 = D\left(x_{(k)} + x_{(k+1)}\right) \approx \frac{a^2}{n}$$

3. Наименьшую дисперсию имеет третья оценка

<u>Примечание:</u> Для получения значения дисперсии для третьей оценки использовали:

Теорема Крамера: Выборочная p-квантиль имеет дисперсию приближенно равную $\frac{1}{n}\frac{p(1-p)}{f^2(x_p)}$, где x_p – истинная p-квантиль, $f(x)$ – плотность распределения наблюдений выборки.

Статистическое сравнение оценок

1. Значение оценок концентрируются в окрестности оцениваемого параметра (свойство несмещенности).
2. С ростом числа наблюдений в выборке точность (величина разброса) оценок улучшается (свойство несмещенности).

То есть размах R и стандартное отклонение S уменьшается.

$$R = \mathbf{max}\, a_i - \mathbf{min}\, a_i \qquad S_a = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(a_i - \overline{a})^2} \qquad \overline{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$$

3. Различные оценки различаются по величине средней ошибки. Откуда следует, что различные способы обработки наблюдений нужно сравнивать по величине среднего значения некоторого критерия качества, например среднего квадрата ошибки.

Виды плотностей распределения основных распределений

Распределение Фишера с m и n степенями свободы

$$f(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} * \left[\frac{m}{n}\right]^{\frac{m}{2}} * x^{\frac{m}{2}-1} * \left[1 + \frac{m}{n} * x\right]^{-\frac{n+m}{2}}$$

Распределение «Хи-квадрат» с n степенями свободы

$$f(x) = \frac{1}{2^{n/2} \cdot \Gamma\left(\frac{n}{2}\right)} * x^{n/2-1} * e^{-x/2}$$

Распределение Стьюдента с m степенями свободы

$$f(x) = \frac{1}{\sqrt{\pi\, m}} \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} * \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}$$

Здесь $0 < x < \infty$, а Г – **гамма**-функция.