Метрология, стандартизация и сертификация

Практика 3

Статистические критерии

(Непараметрические критерии)

Статистические критерии. Введение

Статистический критерий https://ru.wikipedia.org/wiki/Статистический_критерий
строгое математическое правило, по которому принимается или отвергается та или иная статистическая гипотеза с известным уровнем значимости.

Построение критерия представляет собой выбор подходящей функции от результатов наблюдений (ряда эмпирически полученных значений признака), которая служит для выявления меры расхождения между эмпирическими значениями и гипотетическими.

Статистическая гипотеза https://ru.wikipedia.org/wiki/Проверка_статистических_гипотез
предположение о виде распределения и свойствах случайной величины, которое можно подтвердить или опровергнуть применением статистических методов к данным выборки.

Пусть в (статистическом) эксперименте доступна наблюдению случайная величина X, распределение которой P полностью или частично неизвестно. Тогда любое утверждение, относительно , называется статистической гипотезой.

На практике обычно требуется проверить какую-то конкретную и как правило простую гипотезу H_0. Такую гипотезу принято называть нулевой. При этом параллельно рассматривается противоречащая ей гипотеза H_1, называемая конкурирующей или альтернативной.

Этапы проверки статистических гипотез

1. Формулировка основной гипотезы H_0 и конкурирующей гипотезы H_1.

2. Задание уровень значимости α, на котором в дальнейшем и будет сделан вывод о справедливости гипотезы. Он равен вероятности допустить ошибку первого рода.

3. Расчёт статистики φ критерия такой, что:
- её величина зависит от исходной выборки
- по её значению можно делать выводы об истинности гипотезы H_0;
- φ, как функция случайной величины X, также является случайной величиной и подчиняется какому-то закону распределения.

4. Построение критической области. Из области значений φ выделяется подмножество C таких значений, по которым можно судить о существенных расхождениях с предположением. Его размер выбирается таким образом, чтобы выполнялось равенство $P(\varphi \in C) = \alpha$. Это множество C и называется критической областью.

5.Вывод об истинности гипотезы. Наблюдаемые значения выборки подставляются в статистику φ и по попаданию (или непопаданию) в критическую область C выносится решение об отвержении (или принятии) выдвинутой гипотезы H_0.

Ошибки первого и второго рода https://ru.wikipedia.org/wiki/Ошибки_первого_и_второго_рода

		Верная гипотеза	
		H_0	H_1
Результат применения критерия	H_0	H_0 верно принята	H_0 неверно принята (Ошибка *второго* рода)
	H_1	H_0 неверно отвергнута (Ошибка *первого* рода)	H_0 верно отвергнута

Ошибки первого и второго рода являются взаимно-симметричными, то есть если поменять местами гипотезы H_0 и H_1, то ошибки первого рода превратятся в ошибки второго рода и наоборот.
Принято считать, что нулевая гипотеза H_0 соответствует состоянию «по умолчанию» (естественному, наиболее ожидаемому положению вещей). Альтернативная гипотеза H_1 обозначает противоположную ситуацию, которая обычно трактуется как менее вероятная, неординарная.

Ошибку первого рода часто называют ложной тревогой, ложным срабатыванием или ложноположительным срабатыванием — например, анализ показал наличие вещества, хотя на самом его нет.

Ошибку второго рода иногда называют пропуском события или ложноотрицательным срабатыванием — вещество есть, но анализ этого не показал.

Вероятность ошибки первого рода при проверке статистических гипотез называют уровнем значимости и обычно обозначают греческой буквой α.

В статистических тестах обычно приходится идти на компромисс между приемлемым уровнем ошибок первого и второго рода. Зачастую для принятия решения используется пороговое значение, которое может варьироваться с целью сделать тест более строгим или, наоборот, более мягким. Этим пороговым значением является уровень значимости, которым задаются при проверке статистических гипотез. Например, в случае металлодетектора повышение чувствительности прибора приведёт к увеличению риска ошибки первого рода (ложная тревога), а понижение чувствительности — к увеличению риска ошибки второго рода (пропуск запрещённого предмета).

Статистическая значимость

https://ru.wikipedia.org/wiki/Статистическая_значимость

В статистике величину называют статистически значимой, если мала вероятность её случайного возникновения или ещё более крайних величин. Здесь под крайностью понимается степень отклонения тестовой статистики от нуль-гипотезы.

Уровень α значимости теста — вероятность отклонить гипотезу H, если на самом деле она верна (решение известное как ошибка первого рода, или ложноположительное решение).
Популярными уровнями значимости являются 10 %, 5 %, 1 %, и 0,1 %.
Различные значения α -уровня имеют свои достоинства и недостатки.
Меньшие α -уровни дают большую уверенность в том, что уже установленная альтернативная гипотеза значима, но при этом есть больший риск не отвергнуть ложную нулевую гипотезу (ошибка второго рода, или «ложноотрицательное решение»).
Выбор α -уровня неизбежно требует компромисса между следовательно между вероятностями ошибок первого и второго рода.

Виды статистических критериев https://ru.wikipedia.org/wiki/Статистический_критерий

Критерии значимости. Проверка на значимость предполагает проверку гипотезы о численных значениях известного закона распределения:
H_0: $a=a_0$ — нулевая гипотеза. H_1: $a>a0$ ($a<a_0$) или $a\neq 0$ — конкурирующая гипотеза.

Критерии согласия. Проверка на согласие подразумевает проверку предположения о том, что исследуемая случайная величина подчиняется предполагаемому закону. Критерии согласия можно также воспринимать, как критерии значимости. Критериями согласия являются:
Критерий Пирсона
Критерий Колмогорова

Критерии проверки на однородность. При проверке на однородность случайные величины исследуются на факт значимости различия их законов распределения (т.е. проверки того, подчиняются ли эти величины одному и тому же закону).

Это разделение условно, и один и тот же критерий может быть использован в разных качествах.

- - -

Непараметрические критерии - группа статистических критериев, которые не включают в расчёт параметры вероятностного распределения и основаны на оперировании частотами или рангами.
[Критерий Уилкоксона, **Критерий Пирсона (или критерий согласия χ^2 (Хи-квадрат)** ,
Критерий Колмогорова-Смирнова]

Непараметрические критерии - группа статистических критериев, которые включают в расчёт параметры вероятностного распределения (средние и дисперсии).
[**t-критерий Стьюдента, Критерий Фишера**]

КРИТЕРИЙ χ^2 (Критерий согласия Пирсона)

На практике часто приходится встречаться с задачами следующего рода.
Некоторое испытание производится несколько раз, причем известна теоретическая частота появления некоторого события при этом испытании.

Однако на практике фактическая частота оказалась несколько отличной от теоретической.
Надо установить, можно ли объяснить имевшее место расхождение между частотами случайными причинами

Пусть, например, известно, что процент брака при выпуске некоторой продукции в среднем равен a.
Если среди выпускаемой продукции процент брака окажется равным $b \neq a$, то можно ли это расхождение объяснить случайными причинами или это расхождение существенно и вызвано улучшением или ухудшением технологии производства этой продукции?

Для решения этого вопроса поступают следующим образом.
Построим некоторую величину, которую можно было бы принять за меру расхождения между фактической и теоретической частотой. Эту величину называют критерием значимости.

Определим вероятность того, что в силу случайных величин, критерий значимости примет значения, равные или большие того значения, которое получено из опыта.

Если эта вероятность окажется малой, следовательно, достигнуто это значения не в силу случайных причин и расхождение между теоретической и фактической частотами существенно.
(Малая вероятность - расхождение неслучайно)

Если же эта вероятность окажется большой, то расхождение между теоретической и фактической частотами следует признать случайным.
(Вероятность большая - расхождение случайно)

Вопрос о том, какую вероятность нужно считать малой, не может быть решен методами математики; он зависит от характера рассматриваемой задачи. Часто при решении подобных вопросов вероятность считают малой, если она меньше 0,05 (пятипроцентный уровень значимости). Однако следует иметь в виду, что при этом в одном из каждых 20 случаев мы будем утверждать наличие эффекта, не существующего в действительности.

Если такой процент ошибки считается слишком большим, следует принять более высокий, например, 1%-й уровень значимости.

Если нам нужно определить, случайно ли отличается частота появления некоторого события от ожидаемого значения, то применяют так называемый χ^2 критерий.

За меру расхождения между теоретической и наблюдаемой частотой принимают число:

$$\chi^2 = \sum \frac{(\Phi - E)^2}{E}$$

где Φ — фактически полученное значение частоты; E — ожидаемая частота.
Суммирование производится по всем исходам опыта.

Не рекомендуется применять критерий χ^2 в тех случаях, когда какое-либо E меньше 5.
После того, как величина χ^2 найдена, нужно из таблицы величины χ^2 определить вероятность того, что в силу случайных причин χ^2 примет значение, равное или большее того, которое найдено из опыта.

Таблица функции χ^2 составлена по двум аргументам. Одним из них является вероятность *p,* а другим — так называемое «число степеней свободы». Под числом степеней свободы понимают число классов, значения которых можно задать произвольно.
Иными словами, это есть общее число классов минус число ограничений, наложенных на изучаемую систему.

Если найденная по таблице вероятность мала (например, меньше 0,5 или меньше 0,1), то расхождение между опытной и теоретической частотами неслучайно.

Критерий χ^2 применяется также и при сравнении опытных частот, полученных в результате нескольких опытов.

Критерий χ^2

Число степеней свободы	$p=0{,}99$	0,98	0,95	0,90	0,80	0,20	0,10	0,05	0,02	0,01
1	0,00015	0,0006	0,0039	0,015	0,064	1,64	2,70	3,84	5,41	6,63
2	0,02	0,4	0,10	0,21	0,44	3,21	4,60	5,99	7,82	9,21
3	0,11	0,18	0,35	0,58	1,00	4,64	6,25	7,81	9,83	11,34
4	0,29	0,42	0,71	1,06	1,64	5,98	7,77	9,48	11,66	13,27
5	0,55	0,75	1,14	1,61	2,34	7,28	9,23	11,07	13,38	15,08
6	0,87	1,13	1,63	2,20	3,07	8,55	10,64	12,59	15,03	16,81
7	1,23	1,56	2,16	2,83	3,82	9,80	12,01	14,06	16,62	18,47
8	1,64	2,03	2,73	3,49	4,59	11,03	13,36	15,50	18,16	20,09
9	2,08	2,53	3,32	4,16	5,38	12,24	14,68	16,91	19,67	21,66
10	2,55	3,05	3,94	4,86	6,17	13,44	15,98	18,30	21,16	23,20
11	3,05	3,60	4,57	5,57	6,98	14,63	17,27	19,67	22,61	24,72
12	3,57	4,17	5,22	6,30	7,80	15,81	18,54	21,02	24,05	26,21
13	4,10	4,76	5,89	7,04	8,63	16,98	19,81	22,36	25,47	27,68
14	4,66	5,36	6,57	7,79	9,46	18,15	21,06	23,68	26,87	29,14
15	5,22	5,98	7,26	8,54	10,30	19,31	22,30	24,99	28,25	30,57
16	5,81	6,61	7,96	9,31	11,15	20,46	23,54	26,29	29,63	32,00
17	6,40	7,25	8,67	10,08	12,00	21,61	24,76	27,58	30,99	33,40
18	7,01	7,90	9,39	10,86	12,85	22,76	25,98	28,86	32,34	34,80
19	7,63	8,56	10,11	11,65	13,71	23,90	27,20	30,14	33,68	36,19
20	8,26	9,23	10,85	12,44	14,57	25,03	28,41	31,41	35,02	37,56
21	8,89	9,91	11,59	13,24	15,44	26,17	29,61	32,67	36,34	38,93
22	9,54	10,60	12,33	14,04	16,31	27,30	30,81	33,92	37,65	40,28
23	10,19	11,29	13,09	14,84	17,18	28,42	32,00	35,17	38,96	41,63
24	10,85	11,99	13,84	15,65	18,06	29,55	33,19	36,41	40,27	42,98
25	11,52	12,69	14,61	16,47	18,94	30,67	34,38	37,65	41,56	44,31
26	12,19	13,40	15,37	17,29	19,82	31,79	35,56	38,88	42,85	45,64
27	12,87	14,12	16,15	18,11	20,70	32,91	36,74	40,11	44,14	46,96
28	13,56	14,84	16,92	18,93	21,58	34,02	37,91	41,33	45,41	48,27
29	14,25	15,57	17,70	19,76	22,47	35,13	39,08	42,55	46,69	49,58
30	14,95	16,30	18,49	20,59	23,36	36,25	40,25	43,77	47,96	50,89

Вероятности p
в долях 1.
Для получения в %
умножить на 100.

Создать файл в Excel: *Фамилия*_МСС_Пр03

Задача 1

Прочность предмета испытывается с помощью стального шарика, который падает на предмет с определенной высоты.

В одной серии испытанию подвергались 10 предметов, причем 5 из них были разбиты, в другой серии из 20 предметов другого состава не выдержали испытания 14.

Можно ли считать, что прочность предмета обусловливается различием состава или же различие в результатах опытов по сериям объясняется случайными причинами?

Подсказка

0. В расчетах дробных величин сохранять 2 знака после запятой

1. Вычислить общее (по обоим сериям) число разбитых (R) и целых предметов (Z).

$$R = R_1 + R_2, \quad Z = Z_1 + Z_2,$$

где

R_1 - число разбитых предметов в первой серии, Z_1 - целых предметов в первой серии

R_2 - число разбитых предметов во второй серии, Z_2 - целых предметов во второй серии

2. Рассчитать общее (по обоим сериям) теоретические вероятности для предметов P_R быть разбитыми или P_Z уцелеть как

$$P_R = R/(R+Z), \quad P_Z = Z/(R+Z),$$

3. Найти теоретическое число разбитых предметов в 1 и 2 сериях

$$N_{R1} = P_R * N_1, \qquad N_{R2} = P_R * N_2,$$

(N_1 - число предметов в первой серии, N_2 - число предметов во второй серии)

4. Найти теоретическое число целых предметов в 1 и 2 сериях

$$N_{Z1} = P_Z * N_1, \qquad N_{Z2} = P_Z * N_2$$

4. Найти критерий χ^2

$$\chi^2 = \sum \frac{(\Phi - E)^2}{E}$$

Ф — фактически полученное значение частоты (число предметов);
E — ожидаемая частота (число предметов).

Суммирование производится по всем исходам опыта:
всего 4 исхода опыта:
1) разбитое в первой серии, 2) уцелевшее в первой серии,
3) разбитое во второй серии, 2) уцелевшее во второй серии

5. Определить число степеней свободы
- Число степеней свободы находится как число классов, значения которых могут быть заданы произвольно. В нашем случае имеем всего два класса — класс разбитых и класс уцелевших предметов; однако произвольно мы можем задать численность только одного из них, а численность второго класса вычисляется.

6. Обращаясь к таблице «Критерий χ^2» для нужного числа степеней свободы, находим, что какова вероятность нашей гипотезы.
- Если эта вероятность мала, то можно утверждать, что здесь имеется различие свойств испытываемого состава предмета.
- Если эта вероятность не мала, то нельзя утверждать, что здесь имеется различие свойств испытываемого состава предмета.

Задача 2

В первой выборке в N_1 изделий было R_1 случаев брака; в другой выборке в N_2 изделий было R_2 случаев брака. Оправдывается ли с достаточной уверенностью предположение, что совокупности, представляемые этими двумя партиями, различны?

Подсказка

Решать аналогично задаче 1. Данные брать в **МСС_Пр03_Распределение (...).xls**

Задача 3

Выпуск сверхплановой продукции по сменам (в условных единицах): A, B, C
Можно ли считать расхождение в сверхплановом выпуске по сменам случайным?

Подсказка

Найти средний выпуск по сменам E

$$\chi^2 = \sum \frac{(\Phi - E)^2}{E}$$

Ф — фактически полученное значение частоты (число предметов);
E — ожидаемая частота (число предметов).

Данные брать в **МСС_Пр03_Распределение (...).xls**

Задача 4

В таблице (см. **МСС_Пр03_Распределение (...).xls**) приведены данные о процессе работы машины. Испытывались четыре различных режима работы.
Проверить гипотезу, что частота поломок не зависит от метода загрузки

Подсказка
1) Вычислить общее количество циклов и общее количество поломок
2) Вычислить общую теоретическую (ожидаемую) частоту поломок по всем циклам (число поломок на цикл)
3) Вычислить ожидаемое число поломок в каждом цикле
4) Вычислить на основании полученных данных χ^2
5) Определить число степеней свободы
6) По таблице определить для найденного числа степеней свободы и χ^2 вероятность.
7) Записать вывод: есть ли связь режима работы и поломок

Задача 5

При отсчетах по шкалам измерительных приборов последние цифры показаний обычно оцениваются лишь приблизительно в долях деления шкалы.
При этом часто можно отметить предпочтение, которое даже опытные наблюдатели оказывают одним цифрам перед другими.
В таблице (см. **МСС_Пр03_Распределение (...).xls**) приведено распределение NNN случаев оценки последней цифры одним из наблюдателей при отсчете по измерительному прибору в долях деления шкалы.
Имеем мы здесь дело с систематической ошибкой в отсчете или нет?

Подсказка

1) Найти теоретическую вероятность появления цифр 0-9

2) Решать аналогично задачам 3 и 4

Задача 6

Вычислить Таблицу «Критерий χ^2» с использованием функции Excel   ХИ2ОБР

Подсказка

1) Создать новый лист Excel, назвать **Таблица**

2) Ввести в первую строку (начиная со второй ячейки) значения вероятностей (согласно таблице): 0.99, 0.98, … 0.01

3) Ввести в первый столбец (начиная со второй ячейки) число степеней свободы (согласно таблице): 1, 2, … 30

4) В ячейках таблицы, на ввести функцию ХИ2ОБР (Вероятность; Степени_свободы)

5) Отформатировать результаты по числу знаков после запятой согласно таблице «Критерий χ^2»