(MQXFA) Magnet Training and CLIQ

S. Stoynev[1], G. Ambrosio[1], K. Amm[2], J. DiMarco[1], S. Feher[1], P. Ferracin[3], V. Marinozzi[1], S. Prestemon[3], A. B. Yahia[2]

[1]FNAL, [2]BNL, [3]LBNL

MDP meeting

17 January 2024

Based on material presented at MT28 (with authors listed above) and other public material

https://inspirehep.net/files/20a8b705121db69dea138f4808e4b5e7

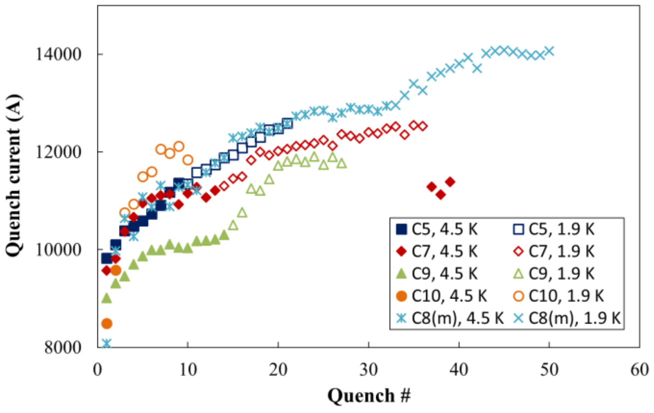
Outline

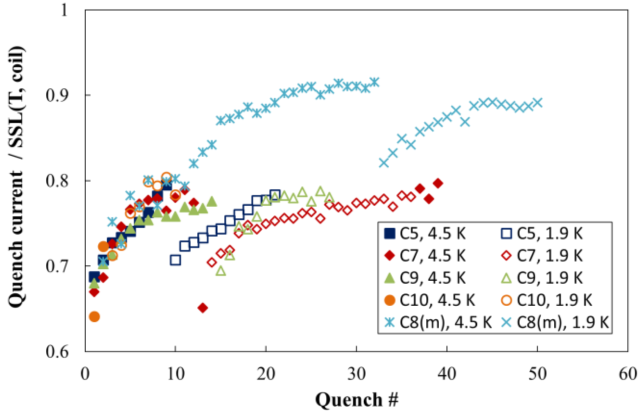
- **Training with CLIQ – initial expectations; findings from an AUP magnets analysis**
- **CLIQ – PS circuit, connections (in MQXFA), “projected” coil training**
- **QXFA coil training “compressed” data, observations and features**
- **Very-very-very brief analysis tools overview (see back up)**
- **Hypotheses testing and results**
- **Reversing CLIQ polarity (consecutively)**
- **Conclusions**

Based on material presented at MT28
and other public material

Training observations and consequences

I : Coils in magnets (can) train independently



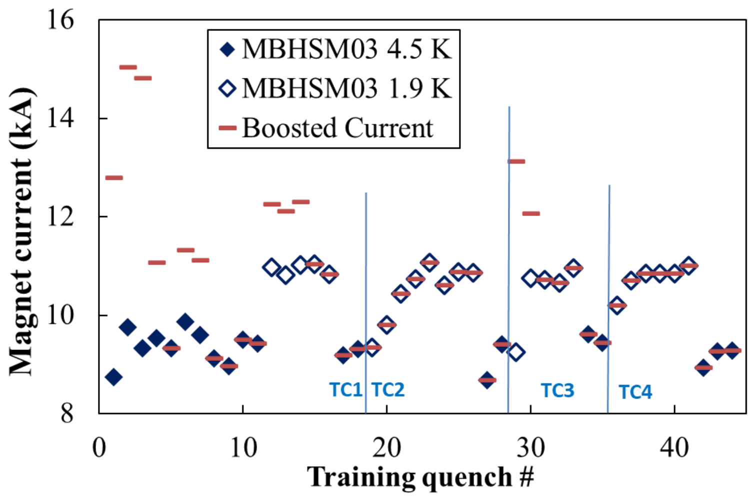


IEEE TRANSACTIONS ON APPLIED SUPERCONDUCTIVITY, VOL. 29, NO. 5, AUGUST 2019 4001206

Analysis of Nb_3Sn Accelerator Magnet Training

Stoyan Stoynev, Kevin Riemer, Alexander V. Zlobin, Giorgio Ambrosio, Paolo Ferracin, GianLuca Sabbi, and Peter Wanderer

II : Short magnet "over-current", which is current above quench current, (can) eliminates training



IEEE TRANSACTIONS ON APPLIED SUPERCONDUCTIVITY, VOL. 33, NO. 5, AUGUST 2023 4702406

Commissioning, Performance, and Effect of the Quench Current-Boosting Device on a Dedicated Superconducting Magnet

S. Stoynev, M. Baldini, and S. Feher

→ **Likely consequence:**

CLIQ will eliminate training in coils with over-current, within some limitations/conditions

The most obvious condition is the level of over-current and duration of it.
It happens that both QCD (left) and CLIQ have typically ~ 30 ms of over-current

(TEXT IN BLUE ON SLIDES IS INFORMATIVE , IGNORE IT IF TOO OVRWHELMING, useful for off-line reading)

Main results from analysis of AUP magnets

- **Coils with over-current do not train <u>at all</u>, ones with under-current do train : only if first magnet quenches are NOT counted (and only for "non-weak" coils)**
  - **In case the first magnet quenches are counted the statistical significance of this difference decreases by a lot**
  - **The statement <u>breaks</u> if very high quench current gain is considered (of the order of the over-current)**

The above conditions can be all linked to CLIQ and explained by accepting that CLIQ reduces (or more precisely – eliminates within limits) coil-training in magnets by over-current induction above quench current.

As a by-product this also proves that coils in magnets (can) train independently.

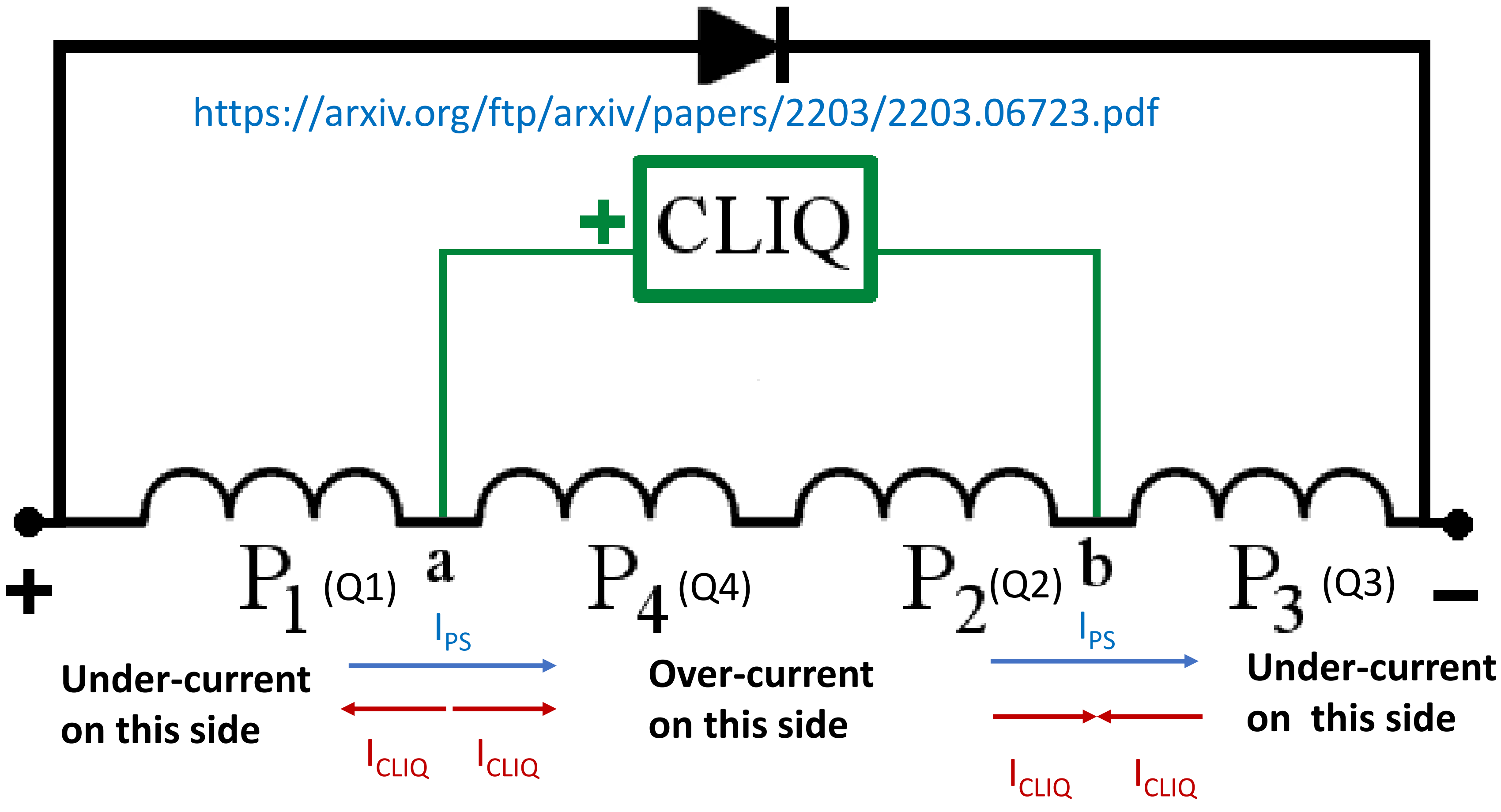
Those and supporting results will be more thoroughly deduced and discussed on the following slides

Article title: Effect of CLIQ on training of HL-LHC quadrupole magnets
Journal acronym: TASC
Article DOI: 10.1109/TASC.2023.3341871
Manuscript Number: MT28-4PoM03-09

- **Journal**: IEEE Transactions on Applied Superconductivity
- **Publication Date**: AUGUST 2024
- **Volume**: 34, **Issue**: 5
- **On Page(s)**: 1-6
- **Print ISSN**: 1051-8223
- **Online ISSN**: 1558-2515
- **Digital Object Identifier**: 10.1109/TASC.2023.3341871

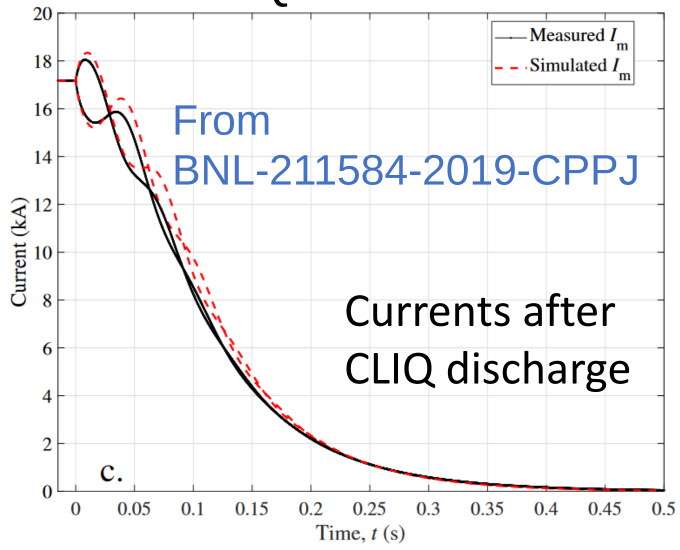
MQXFA connections and quadrants (Q)

MQXFA electrical order of the coils and position of the CLIQ terminals



(this is an oversimplified schematics)

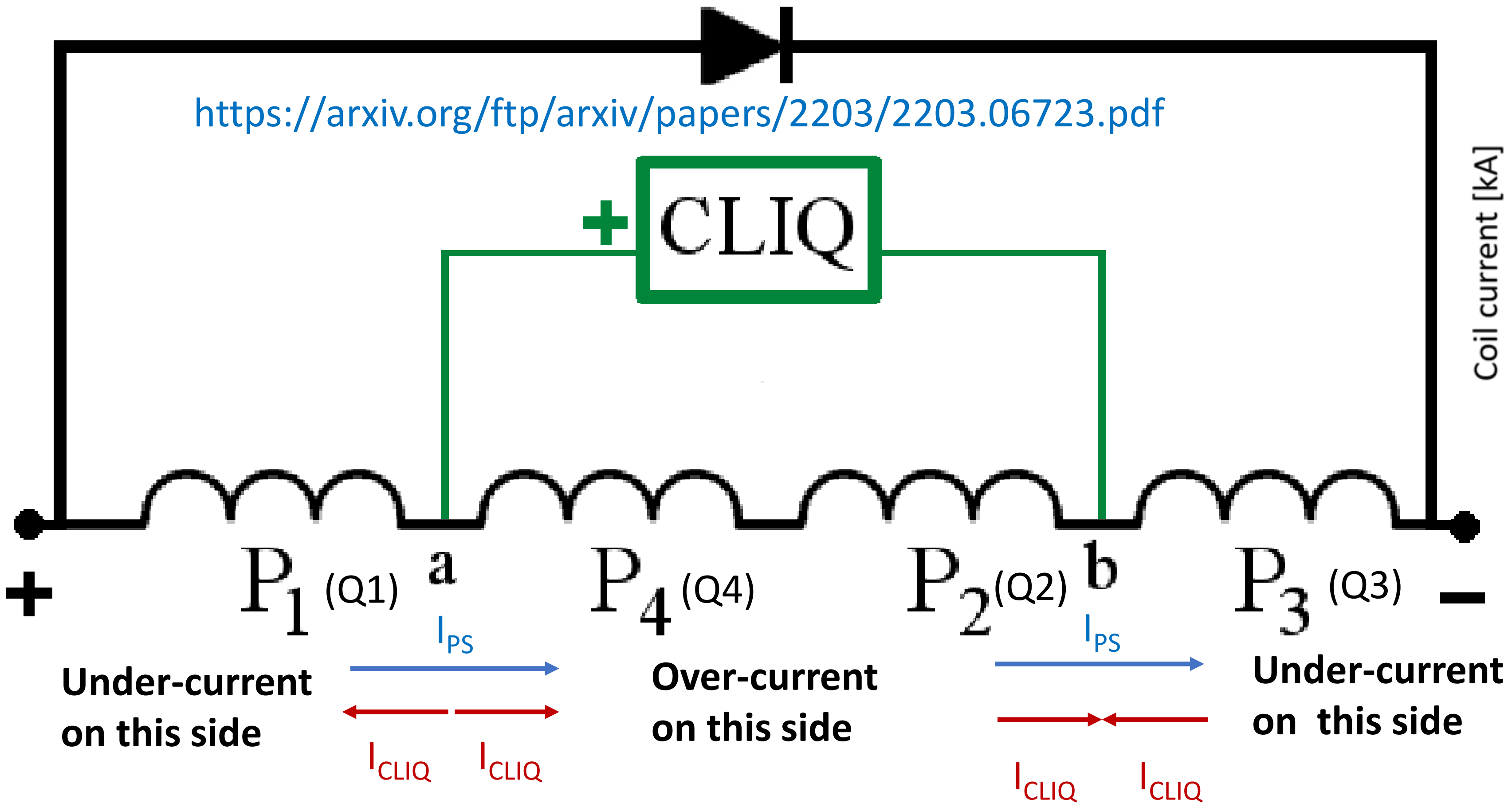
MQXFAP1



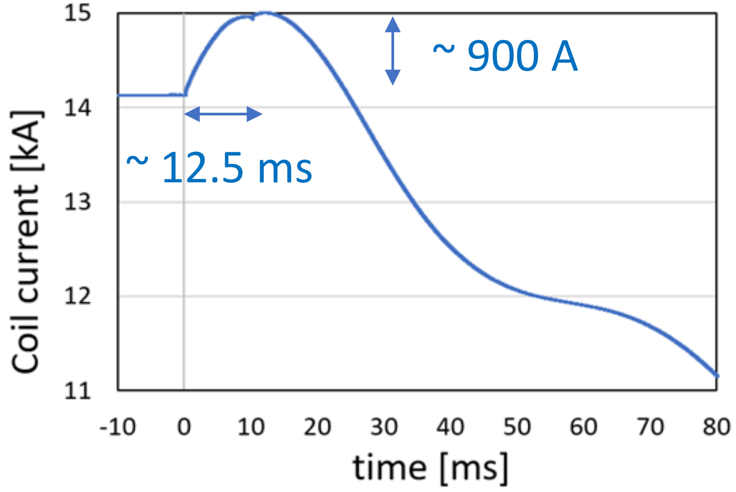
Under-current (in Q1 and Q3) and over-current (in Q2 and Q4) is ~1 kA, note the time scale too. Duration is ~ 20-30 ms with a peak/extremum at ~ 15 ms.

MQXFA connections and quadrants (Q)

MQXFA electrical order of the coils and position of the CLIQ terminals



MQXFA13
(quench one)



Those numbers will somewhat depend on quench current

900 A = 3 x 300 A

300 A is a good unit to use, it is the difference between the acceptance current (16 530 A) and the nominal current (16 230 A); clearly, we think 300 A is a safe margin for quenches at nominal current level

MQXFA training with CLIQ – data "tiers"

- **If the "over-current" mechanism for training works with CLIQ one expects:**
  - **No Q2/Q4 quenches with low current gain (<300 A? <600 A? <900 A?) for "non-weak" coils**
  - **<u>Statistically</u> – more Q1/Q3 training coils than Q2/Q4 training coils**
  - **<u>Statistically</u> – more Q1/Q3 quenches than Q2/Q4 quenches** (somewhat depends on test plans)

Let's define INCLUSIVE "tiers" of data:

- Coil performance: "non-weak" coils (no detraining of >|300 A|), "all" coils
- "Quenched" coils: current gain "< 300 A", "< 600 A", "no limits" (same as "<900 A")

<u>This is the parameter space</u>: **(coil condition, quench current gain condition)**

The "300 A" (unit) for quench current gain/detraining is what the difference between "nominal" and "acceptance" current for MQXFA magnets is.

Let's count quenches in coils (individually) at 1.9 K up to nominal current (+ 10 A) – all magnets reached it
And all magnets to MQXFA13 were pre-stressed the same way

Next slides shows training curves and a "compressed" version of data following categorization based on the above

QXFA coil training to nominal current (+10 A) at 1.9 K

A: Training quenches with gain below 300 A,
B: Training quenches with gain below 600 A,
C: All training quenches
D: Detraining quenches with |ΔI| > 300 A (counted from previous quench)

The very first quench/high current trip in a magnet is removed from all considerations – **not counted** (no use in the context)
"Gain" is calculated with respect to the maximal power supply current provided so far to the magnet

1.9 K

Coil Order: Q1|Q2|Q3|Q4

C1 = FNAL coil
C2 = BNL coil

(shaded cell with 1) -- **"Weak" coils**

	MQXFA03				MQXFA04				MQXFA05			
	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1
A/B/C :	0/0/0	0/0/0	0/0/0	4/7/8	1/1/1	3/4/4	0/0/0	0/0/0	3/4/4	0/0/1	1/1/1	0/0/0
D :	0	0	0	1	0	1	0	0	0	0	0	0
	coil 204	coil 110	coil 202	coil 111	coil 203	coil 113	coil 206	coil 112	coil 207	coil 116	coil 209	coil 115

	MQXFA06				MQXFA07				MQXFA08				(MQXFA08b)
	C1	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	(C2 | Q3)
A/B/C :	1/1/1	0/0/0	0/2/2	0/0/0	1/1/1	0/0/0	6/7/7	0/0/0	0/0/0	0/0/1	0/0/0	0/0/0	(2/2/3)
D :	0	0	0	0	0	0	1	0	0	0	"0"	0	(0)
	coil 122	coil 119	coil 211	coil 123	coil 212	coil 124	coil 214	coil 114	coil 215	coil 126	coil 213	coil 128	coil 219

"0" – degraded / detrained later

	MQXFA10				MQXFA11				MQXFA13			
	C1	C2	C1	C1	C2	C2	C1	C1	C2	C1	C2	C1
A/B/C :	8/8/8	0/0/0	5/6/6	0/0/0	1/1/1	0/0/0	5/6/6	0/0/0	14/14/14	0/0/1	3/4/4	0/0/0
D :	0	0	0	0	0	0	0	0	"0"	0	0	0
	coil 132	coil 221	coil 131	coil 129	coil 223	coil 222	coil 134	coil 135	coil 227	coil 139	coil 229	coil 141

At the end I only care if the number > 0

Observations

- **Only three Q2/Q4 coils quenched with above 600 A gain (once) – those are Q2s in MQXFA05/08/13.**
  - The quench current gains were 891 A, 789 A and 663 A, respectively.
  - All of them were the second training quench in a magnet.

- **There were no other training quenches to nominal current (+ 10 A) in Q2/Q4 (excluding "weak" coils and all "first" training quenches/trips in a magnet)**

- There were only two other training quenches in Q2 (MQXFA10) just above the nominal current, they had gains of +55 A and -5 A ("loss")

- **There were no Q1/Q3 coils with gains above 600 A (a couple were within 50 A).**

The following coils quenched **first in a magnet**:

Q2 in MQXFA03/04/05/06/08 and Q4 in MQXFA07 and never quenched after that.
Q3 quenched first in MQXFA10 and had many training quenches.
Q1 quenched first in MQXFA13 and had many training quenches.

A high-current trip in MQXFA11 occurred before any training quenches.

Coils quenched – in Q-groups: consistency question(s)

Up to and including MQXFA13

Magnets tested: 9 ; coils tested: 37 (including a replaced coil in Q3, in MQXFA08b)

Three coils had detraining quenches of more than 300 A – one Q4 (MQXFA03), one Q2 (MQXFA04) and one Q3 (MQXFA07). Those are **"weak" coils** and are excluded from most of the analysis (such large detraining is not a good sign for any coil). Two other coils detrained/degraded only after reaching nominal current and are not excluded – one in Q2 (MQXFA08) and one in Q1 (MQXFA13). **In total there are 34 "non-weak" coils to analyze.**

Let's count, separately for different Qs, how many coils had ("non-weak", "< 600 A")
at least one training quench to nominal current (+ 10 A), but at below 600 A current gain from the previous max current.

	Q1	Q2	Q3	Q4	Q1+Q3	Q2+Q4
Total count	9	8	9	8	**18**	**16**
Quenched	7	0	6	0	**13**	**0**
Fraction quenched	0.78	0.00	0.67	0.00	**0.72**	**0.00**

The main, but not the only, question is this:

Is the observed fraction for Q2+Q4 (0.00) statistically consistent with the observed fraction for Q1+Q3 (0.72)?

(Note that I don't pose a direct question on what **the "true" fraction** is, I don't necessarily care)

We need quantitative answers, no need to speculate

Statistical analysis (see back up for details and clarifications)

We count coils which quench at least once but according to some rules and grouping. Quenching of a coil is a Bernoulli process, with perceived probability to quench ***p*** and follows a Binomial distribution. However, we don't know what the "true" ***p*** is.

*In Statistics, a **contingency table** provides a way of portraying data that can facilitate calculating probabilities. The table displays sample values in relation to different variables that may **be related or contingent on one another**.*

*All statistical tests have Null Hypothesis. For most tests, the null hypothesis is that there is no relationship between variables of interest or that there is no difference among groups. **H0** is the symbol for it, and it is pronounced H-naught.*

*The Alternative Hypothesis is the logical opposite of the null hypothesis. The acceptance of the alternative hypothesis follows the rejection of the null hypothesis. **H1** is the symbol for it.*

For small samples, and especially for **2 x 2 tables**, there are two well-known "exact tests" for determining the equality of two (binomial) probabilities:

Fisher's exact test and Barnard's exact test
(both use some constraints/assumptions).

One can also use a Log-Likelihood ratio test (G-test) with Yates' correction (for low statistics)

Example of a 2x2 contingency table:

	Q1	Q3
Not quenched	**2**	**3**
Quenched	**7**	**6**
Fraction quenched	0.78	0.67

This is our 2x2 contingency table

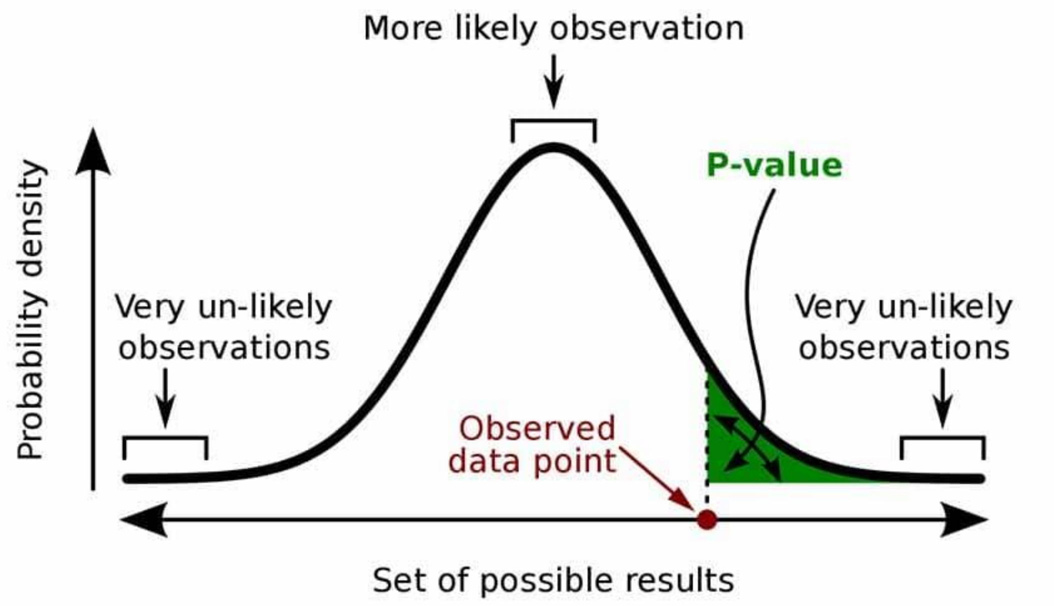
Those tests are available in the "R" Project for Statistical Computing with existing on-line support (and I cross-checked results with known calculations)

https://rdrr.io

P-value and Significance Level

Statistical tests produce a **p-value**, or probability value, which tells how likely it is that certain data could have occurred under the null hypothesis. **It is the probability of the observed outcome or something more extreme than the observed outcome to ocurr, computed under the assumption that the null hypothesis is true.** For instance, if the p-value is 0.05, that means that 5% of the time you would see a test statistic at least as extreme as the one you found if the null hypothesis was true. There is also additional info contained in Confidence Intervals (back up slides).

https://www.simplypsychology.org/p-value.html



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Note that p-values can be one-sided (figure) or two-sided if you count "extreme" probabilities at both tails. A pre-defined **Significance Level (SL)** is the cut on the p-value beyond which we want to reject the Null Hypothesis. It could be **as high as 5% or as low as 10^{-7}** (how "sure" we want to be in rejecting H0).

The p-value can only tell whether the null hypothesis is supported (probabilistically) by measurements. It cannot tell whether the alternative hypothesis is true.

We will use SL = 0.1% for assessment.
(this is approximately the probability to get 10 tails out of 10 flips of a "fair" coin)

Quenching in coils: Q1 vs Q3

Remember: for each test, we only care if a coil quenched at all, we don't care how many times it quenched!

We don't know of anything that will have different effect on training in Q1 vs Q3.
We can test this hypothesis and provide p-values from the tests mentioned.
In this case we have no reason to expect that either Q1 or Q3 should perform better so we should use the two-sided p-values (generally, those are more "conservative" and less biased but can overestimate).

H0 : there is no relation between coils (Q1 vs Q3) and quenching ("< 600 A").

("non-weak", "< 600 A")

	Q1	Q3
Not quenched	**2**	**3**
Quenched	**7**	**6**
Fraction quenched	0.78	0.67

This is our 2x2 contingency table

	P-value for one-sided test	P-value for two-sided test
Fisher's test	**0.500**	**1.00**
Barnard's test	**0.369**	**0.739**
G-test(Yates' corr.)		**1.00**

No grounds to reject H0

We will use SL = 0.1% for assessment.

Quenching in coils: Q1/Q3 vs Q2/Q4, **includes the first quenches for this test only**

Let's look at the data for Q1/Q3 vs Q2/Q4 but include also the first quenches, so if a coil quenched as the first in the magnet we count is as a "quenched" coil
(and we accept this quench as fulfilling any quench current condition we set, like "<600 A").
From the table with quenches we had, and the explicit list of first quenches we'll construct a 2x2 contingency table and set the Null hypothesis in the meantime.

H0 : there is no relation between specific group of coils (Q1/Q3 vs Q2/Q4) and quenching ("< 600A", +1st quenches).

("non-weak", "< 600 A" +1st quenches)

	Q1+Q3	Q2+Q4
Not quenched	5	10
Quenched	13	6

	P-value for one-sided test	P-value for two-sided test
Fisher's test	0.0450	0.0824
Barnard's test	0.0309	0.0583
G-test(Yates' corr.)		0.0892

No grounds to reject H0

Nevertheless, you see some weak indications quenching may not be the same for different Q-groups.
Still not a strong support to claim it.

We will use SL = 0.1% for assessment.

Quenching in coils: Q1/Q3 vs Q2/Q4

Let's look at the same data but exclude the first quenches, i.e. we this is the question we asked for the same data:
is the fraction of quenched coils for Q2+Q4 (0.00) consistent with the fraction of Q1+Q3 (0.72)?

In the context of tables, the question we ask, and asked, is:
are quenching (the rows) and coil groups (the columns) related/associated in some way,
i.e. do proportions between groups differ from each other significantly.

H0 : there is no relation between specific group of coils (Q1/Q3 vs Q2/Q4) and quenching ("< 600A").

("non-weak", "< 600 A")

	Q1+Q3	Q2+Q4
Not quenched	**5**	**16**
Quenched	**13**	**0**

→

	P-value for one-sided test	P-value for two-sided test
Fisher's test	**9.2×10^{-6}**	**9.8×10^{-6}**
Barnard's test	**3.6×10^{-6}**	**7.2×10^{-6}**
G-test(Yates' corr.)		**1.6×10^{-5}**

This is about the same as 25 heads out of 25 flips, is it a "fair-coin", what is your call?

Reject H0

Based on data AND SIGNIFICANCE LEVEL OF 0.1 % we must reject the null hypothesis and accept that there is relation between specific group of coils (Q1/Q3 vs Q2/Q4) and quenching (as defined).

The SL is a choice, but we see p-values change very substantially by just removing the first magnet quenches from consideration.

Quench current gain limits

Let's see what is the role of the quench current gain limit (data on the next slide)

We'll see that by changing the quench current gain to "no limits" (i.e. we include quenches with current gains above 600 A unlike the lower "tier**")** **the statistical results change by orders of magnitude!**

We'll see that there is no much difference between
"<300 A" and "< 600 A" data "tiers" (which just means that ~all quench current gains are under 300 A and we counted them all within "<300 A").

"Weak" coils are <10% of all and have little impact on the analysis (statistically) but clearly make a difference of what statements we can make (if we were to include them).

Quenching in coils: Q1/Q3 vs Q2/Q4 (same H0 but for all different categories)

Data tiers

Data tier			Q1+Q3	Q2+Q4		P-value (1-s)	P-value (2-s)
Non-weak coils (34)	"< 300 A"	Not quenched	6	16	Fisher's test	3.4×10^{-5}	3.7×10^{-5}
		Quenched	12	0	Barnard's test	2.1×10^{-5}	4.1×10^{-5}
					G-test(Yates' corr.)		5.8×10^{-5}
	"< 600 A"	Not quenched	5	16	Fisher's test	9.2×10^{-6}	9.8×10^{-6}
		Quenched	13	0	Barnard's test	3.6×10^{-6}	7.2×10^{-6}
					G-test(Yates' corr.)		1.6×10^{-5}
	"no limits"	Not quenched	5	13	Fisher's test	2.4×10^{-3}	2.6×10^{-3}
		Quenched	13	3	Barnard's test	9.9×10^{-4}	1.9×10^{-3}
					G-test(Yates' corr.)		4.6×10^{-3}
All coils (37)	"< 300 A"	Not quenched	6	16	Fisher's test	4.7×10^{-4}	6.3×10^{-4}
		Quenched	13	2	Barnard's test	1.7×10^{-4}	2.8×10^{-4}
					G-test(Yates' corr.)		8.9×10^{-4}
	"< 600 A"	Not quenched	5	16	Fisher's test	1.4×10^{-4}	1.9×10^{-4}
		Quenched	14	2	Barnard's test	4.8×10^{-5}	8.2×10^{-5}
					G-test(Yates' corr.)		2.7×10^{-5}
	"no limits"	Not quenched	5	13	Fisher's test	6.4×10^{-3}	8.6×10^{-3}
		Quenched	14	5	Barnard's test	3.9×10^{-3}	7.7×10^{-3}
					G-test(Yates' corr.)		1.2×10^{-2}

Coil fabrication origin

H0: there is no relation between quenching of FNAL coils and where they are placed (Q1/Q3 vs Q2/Q4).

(“non-weak”, “< 600 A”)

	C1 in Q1/Q3	C1 not in Q1/Q3
Not quenched	**0**	**14**
Quenched	**4**	**0**

	P-value (1-s)	P-value (2-s)
Fisher’s test	3.3×10^{-4}	3.3×10^{-4}
Barnard’s test	7.2×10^{-5}	7.2×10^{-5}
G-test(Yates’ corr.)		6.1×10^{-4}

Reject H0

We see that the Q1/Q3 vs Q2/Q4 discrimination holds even if we test it on FNAL coils alone.

So, the coil fabrication origin can not be the explanation for the difference observed.

Other sub-sample tests yield statistically non-conclusive results because there is not enough data in them.

C1 = FNAL coil
C2 = BNL coil

Recap

1. p-values change by orders of magnitude by just removing the first magnet quenches from consideration.

2. p-values change by orders of magnitude by changing the upper limit on the quench current gain threshold.

3. Q2/Q4 coils training features are unlike Q1/Q3 coils, and this is not related to their fabrication origin.

Q2 training quench(es) above nominal current

I did say that there should be no quenches in non-weak coils in Q2/Q4 within some current gain constraints and there weren't. However, for one Q2 coil there were two quenches above the nominal current, while training was on-going (seemingly). Why did it quench at all? How do results change if we were to include this case? This was in MQXFA10.

1) There were two quenches in that Q2 coil, one at 16253 A (gain +55 A) and the other at 16525 A (loss -5 A).

Results for this case (one quenched coil in Q2/Q4) is shown on the right.
Quench currents are fairly high, and one can argue those are approaching the coil conductor limit (critical surface; due to possible degradation). Close to critical surface previous training quenches don't matter anymore (the magnet is not in training mode). While this argument may be plausible, another one is stronger.

	Q1+Q3	Q2+Q4
Not quenched	5	15
Quenched	13	1

	P-value (1-s)	P-value (2-s)
Fisher's test	1.0×10^{-4}	1.1×10^{-4}
Barnard's test	3.0×10^{-5}	5.8×10^{-5}
G-test(Yates' corr.)		1.9×10^{-4}

2) In AUP/LARP magnets we see detraining quenches, some with large current loss. In fact, in all the MQXFA magnets (and even earlier models) we tested so far only MQXFA05 had no detraining quench (MQXFA06 had some at 4.5 K, most of the training was there; even MQXFA10 had a detraining quench in a coil beyond Q2). Moreover, Q2 in MQXFA10 had a second quench which was 5 A below the previously reached highest magnet current (clearly a detraining quench).

So, it is not unconceivable that despite the Q2 in MQXFA10 reaching much higher than quench current because of CLIQ, it could still quench at lower than previous quench current.

This is not a "nominal" training behavior of a coil and may signal potential issues with the coil (same for any other "detraining" or at least "weak" coil). The fact it is observed after CLIQ current boosts is the worrisome part for the coil!

"Weak" coils

According to our nomenclature there were three "weak" coils,
and in particular there was **one "weak" coil in each MQXFA03 and MQXFA4**

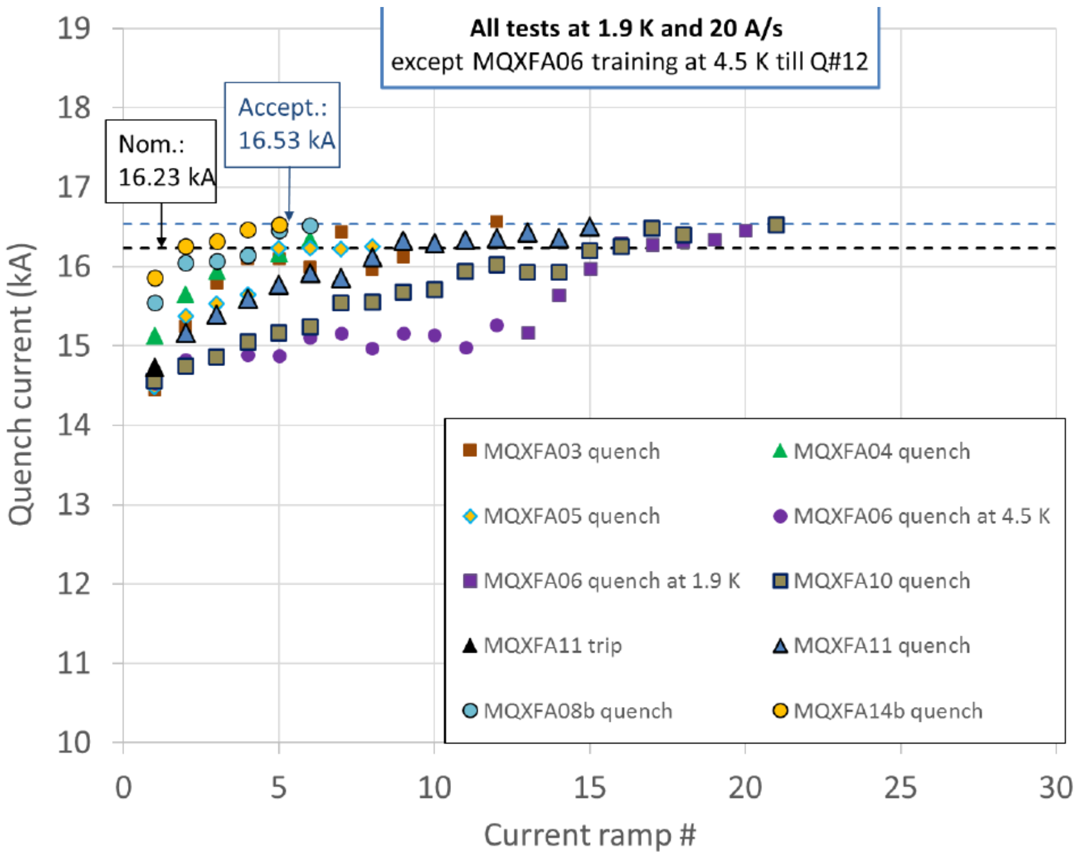
- **"Weak" coils with previous "over-current" suggest huge detraining quenches, could this be a bad omen?**
  According to our expectations and analysis , previous slide, "weak" coils which saw "over-current" mean significant coil detraining (with respect to the reached over-current or conductor limit). If a coil saw detraining quenches at high currents, chances are they would happen again at those current levels (and may deteriorate further). Then "weak" coils are indeed what the notation suggests. There is some evidence of that already.

- **Two out of two <u>retested</u> "weak "coils with "over-current" needed (short) retraining**
  As presented in MT28, by Maria, the first MQXFA cryo-assembly contains magnets MQXFA03 and MQXFA04 and it was tested at FNAL. Although both magnets were trained to 16 530 A at BNL and had hours of operations at that level, each of them quenched once below that level during the cryo-assembly testing - at 16525 A/16385 A (no other quenches occurred). Quenches were in the coils we declared "weak" based on their performance (at BNL) with CLIQ: coils 111 (Q4) and 113 (Q2), respectively.

"Weak" coils are what the name suggests – they are likely to perform worse in consequent tests. This is especially relevant for coils with over-current (which serves as a diagnostics tool here) – the "weakness" must be much larger for them.

MQXFA14b

Pre-stress target increased in MQXFA14b – the current level to count quenches to may be better to change



No matter the level, the first quench was in Q2 and then several quenches followed in Q3.

This is in line with the observations in other magnets - no Q2/Q4 training quenches to operational levels after the first magnet quench.

https://inspirehep.net/files/7e69e9f98d892d88c8d982967a183c8c

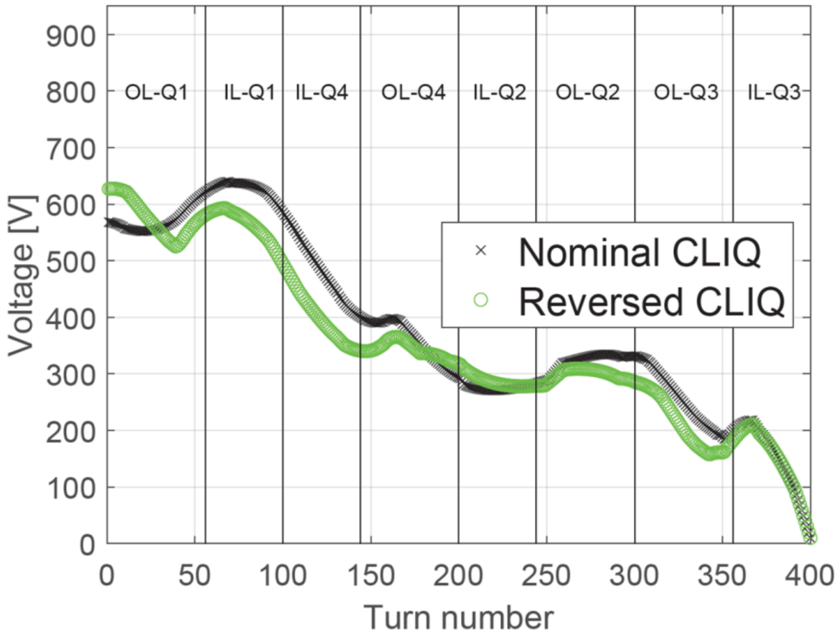
(in private conversations, people from CERN confirmed they also observe non-training behavior of Q2/Q4 in their quadrupole magnets, which are longer but otherwise similar to the AUP-ones)

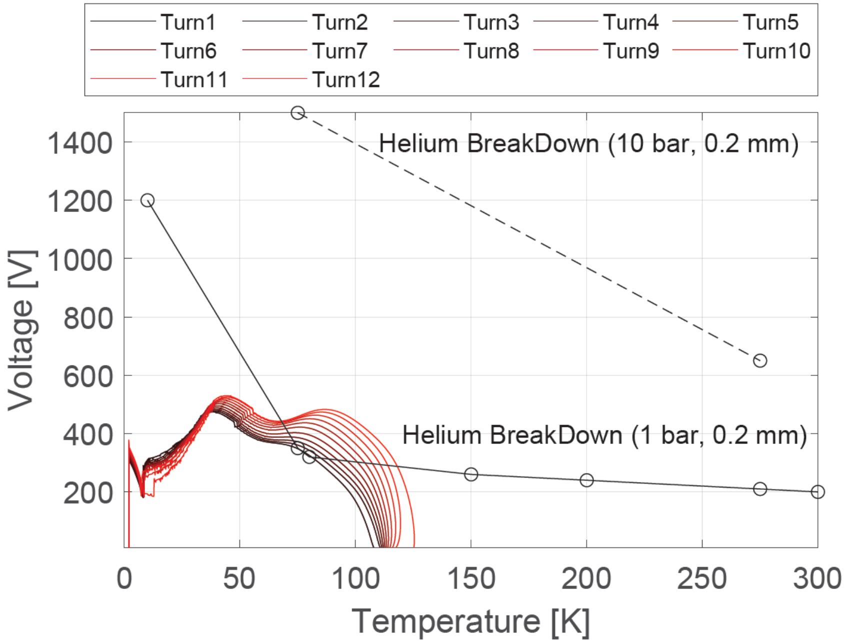
Safety of reversing CLIQ polarity in successive ramps

AUP magnets

SIMULATIONS (STEAM-LEDET):

Vittorio M.

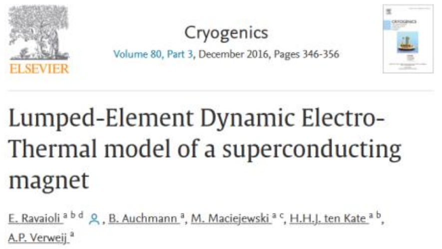




IEEE TRANSACTIONS ON APPLIED SUPERCONDUCTIVITY, VOL. 28, NO. 3, APRIL 2018

STEAM: A Hierarchical Cosimulation Framework for Superconducting Accelerator Magnet Circuits

L. Bortot, B. Auchmann, I. Cortes Garcia, A. M. Fernandez Navarro, M. Maciejewski, M. Mentink, M. Prioli, E. Ravaioli, S. Schps, and A. P. Verweij



Cryogenics

Volume 80, Part 3, December 2016, Pages 346-356

Lumped-Element Dynamic Electro-Thermal model of a superconducting magnet

E. Ravaioli, B. Auchmann, M. Maciejewski, H.H.J. ten Kate, A.P. Verweij

Maximal voltages to ground in coil-turns (outer layers, OL; inner layers, IL) for "nominal" and "inverted" CLIQ polarity.

Coil-turn-to-heater voltage development (worst layer) and breakdown voltage contours in He gas for relevant distance, "nominal" case.

Maximum voltages to ground are about the same with inverted CLIQ polarity.

For a short time, heater-to-coil voltages are significantly higher with "reversed" CLIQ because (heater+CLIQ) polarities were pre-arranged to minimize voltage. Reversing CLIQ polarity adds, instead of subtracts, voltages. However, the insulation is at its weakest point at higher temperature where CLIQ polarity does not matter (CLIQ is discharged by the time this temperature is reached).

Without a full analysis yet, indications are that reversing CLIQ polarity as is, is safe for the AUP magnets.

Conclusions

If we accept the alternatives of rejected H0’s, make reasonable assumptions if no apparent contradictions or known issues exists, and look at trends in data, we could come to the following picture:

- **The persistent difference between Q1/Q3 and Q2/Q4 quenching (regardless of coil origin),**
- **the crucial importance of the exclusion of first quenches (in magnets) for this analysis**
- **and the apparent importance of current gain threshold seen,**

can be all linked to CLIQ and explained by accepting that CLIQ eliminates coil-training in magnets, within some limits, by over-current induction above quench current.

As a by-product this also proves that coils in magnets (can) train independently.

As “expected” (slide3)

Another by-product: we can use CLIQ to identify “weak” coils, if we have to

If we want faster training of MQXFA/B or other magnets
we could develop safe and efficient swapping of CLIQ polarity
(supposedly it needs to be done few times per magnet training session, but it depends on CLIQ energy)

Moreover, no quenches are actually needed for “training”–
the “training” to “acceptance” current can be achieved by series of high-current trips
(and we save detection and validation times - less MIITs)

Back up slides

Back up: Bernoulli Trials

https://www.statisticshowto.com/bernoulli-trials/
https://link.springer.com/referenceworkentry/10.1007/978-0-387-32833-1_27
https://link.springer.com/referenceworkentry/10.1007/978-0-387-32833-1_24

The three assumptions for Bernoulli trials are:

1. Each trial has two possible outcomes: Success or Failure.
2. The **probability** of Success (and of Failure) is constant for each trial; a “Success” is denoted by the letter *p* and “Failure” is $q = 1 - p$.
3. Each trial is independent; The outcome of previous trials has no influence on any subsequent trials.

Bernoulli distribution:

$$P(X = x) = \begin{cases} p & \text{for } \; x = 1 \\ q = 1 - p & \text{for } \; x = 0 \end{cases}$$

Back up: Binomial distribution

https://link.springer.com/referenceworkentry/10.1007/978-0-387-32833-1_34

If X1, … , Xn are *n* independent **random variables** following a **Bernoulli distribution** with a **parameter** *p*, then the random **variable** X = X1 + … Xn follows a binomial distribution *B*(*n*, *p*).

Technically, the Bernoulli distribution is a Binomial distribution with n = 1.

Back up: changing probabilities in trials

If probabilities in Bernoulli trials are not the same, then the variance of this distribution (sum of Binomials with different probabilities) is always smaller than or equal to the variance of the Binomial distribution with the mean value of probabilities.

Sample variance:

$$\sigma_p^2 = \frac{1}{n}\sum_{i=1}^{n}(p_i - \overline{p})^2$$

with sample mean:

$$\overline{p} = \frac{1}{n}\sum_{i=1}^{n}(p_i)$$

Then

$$\sigma_p^2 = \frac{1}{n}\sum(p_i^2) + \frac{n}{n}\overline{p}^2 - \frac{1}{n}\sum 2(p_i\overline{p}) = \frac{1}{n}\sum(p_i^2) + \overline{p}^2 - 2\overline{p}^2 = \frac{1}{n}\sum(p_i^2) - \overline{p}^2$$

Thus

$$\sum(p_i^2) = n(\sigma_p^2 + \overline{p}^2)$$

Bernoulli variance:

$$Var(p_i) = p_i(1 - p_i)$$

$$Var\left(\sum p_i\right) = \sum[p_i(1 - p_i)] = \sum(p_i) - \sum(p_i^2) = n\overline{p} - n(\sigma_p^2 + \overline{p}^2) = n\overline{p}(1 - \overline{p}) - n\sigma_p^2 = Var(\overline{p}) - n\sigma_p^2$$

≥ 0

Reminder – Bernoulli trials and Binomial distributions

Any event that has exactly two outcomes with a fixed probability is called a Bernoulli random variable. Every Bernoulli distribution has a probability, *p*, describing the probability of that event occurring.

The Binomial Distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent events. In other words, the Binomial Distribution is the sum of n independent Bernoulli random variables.

For a single trial (n=1), the binomial distribution is a Bernoulli distribution.

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

When we run *n* independent experiments, each having a Bernoulli distribution with parameter *p,* and we want to know the probability of having x successes, the probability function will be:

$$f(x) = \begin{cases} \binom{n}{x} p^x * (1-p)^{n-x} & if\ x = 0,1,..,n \\ 0 & otherwise \end{cases}$$

Such a random variable is expressed as *X ~ B(n,p)*

For all binomial distributions B(n, p) where n is the number of trials and p is the probability of success:

$$EV_{binomial} = np$$ Expected Value, E

$$SE_{binomial} = \sqrt{np(1-p)}$$ Standard Error (SE^2 = Variance, V)

$$E(X) = \sum_x x * f(x)$$

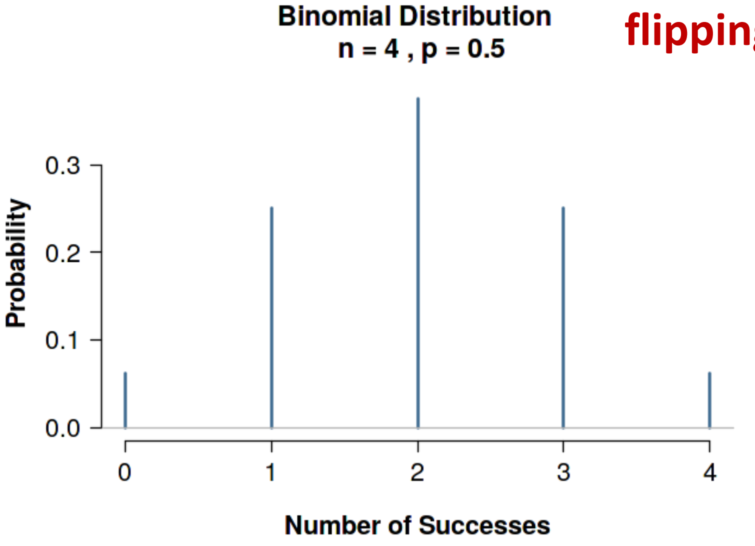
$$V(X) = \sum_x (x-\mu)^2 * f(x) = E(X^2) - E^2(X)$$

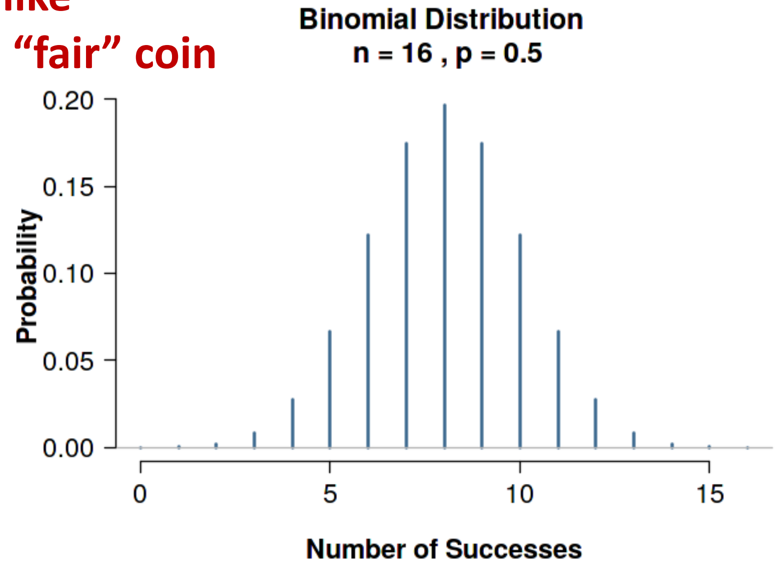
https://towardsdatascience.com/understanding-bernoulli-and-binomial-distributions-a1eef4e0da8f

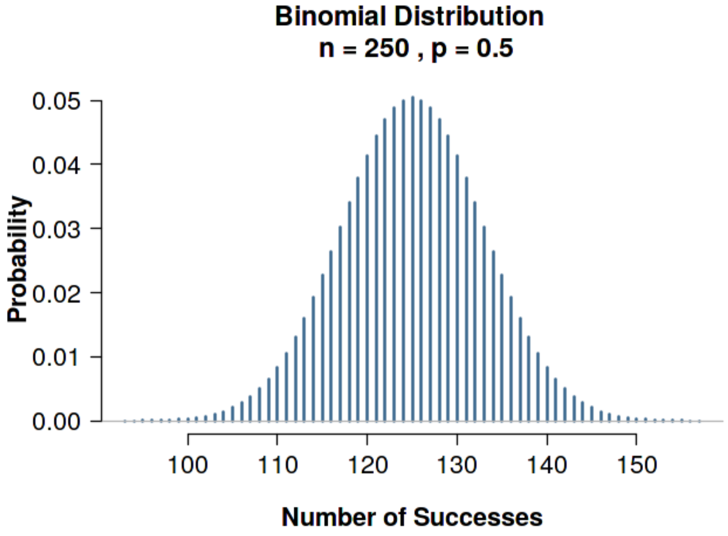
https://discovery.cs.illinois.edu/learn/Polling-Confidence-Intervals-and-Hypothesis-Testing/Bernoulli--Binomial-Random-Variables/

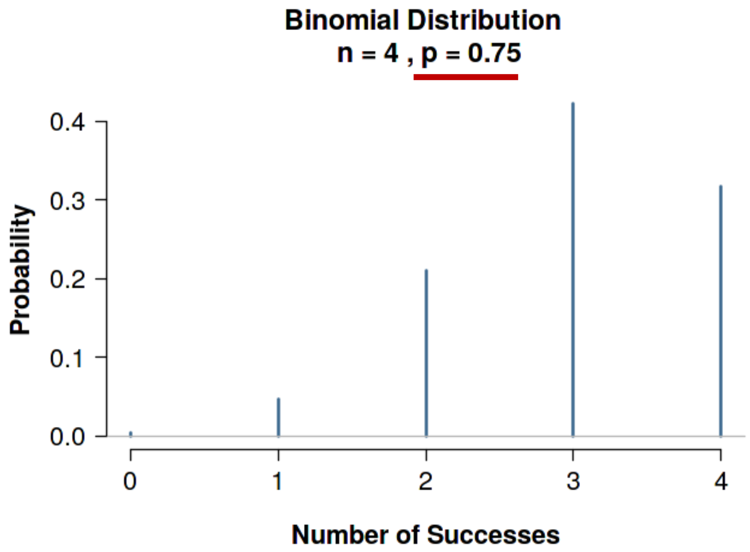
Reminder – example of Binomial distributions

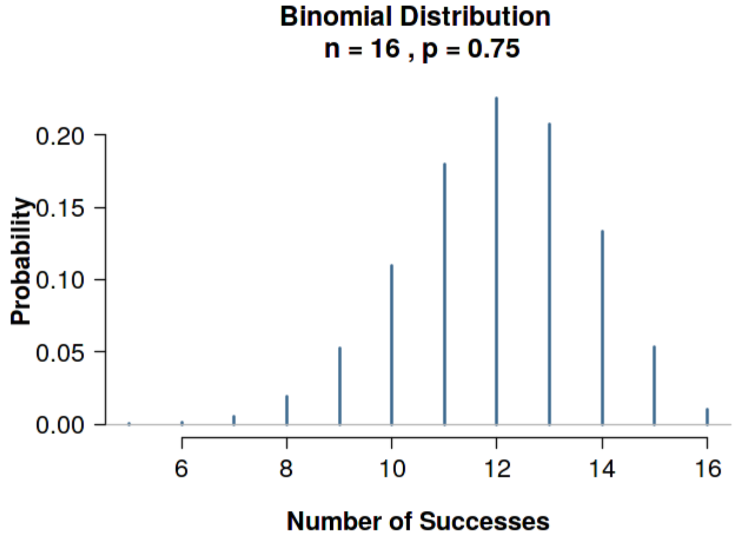
p = 0.5 is like flipping a "fair" coin

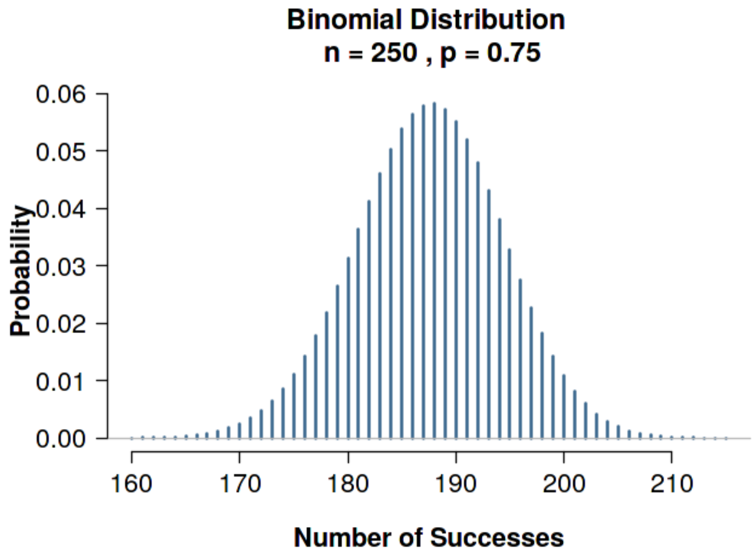












https://shiny.rit.albany.edu/stat/binomial/

Reminder – hypothesis testing

Hypothesis Testing is a type of statistical analysis in which you put your assumptions about a population parameter to the test. It is used to estimate the relationship between two statistical variables.

All statistical tests have Null Hypothesis. For most tests, the null hypothesis is that there is no relationship between variables of interest or that there is no difference among groups. **H0** is the symbol for it, and it is pronounced H-naught.

The Alternative Hypothesis is the logical opposite of the null hypothesis. The acceptance of the alternative hypothesis follows the rejection of the null hypothesis. **H1** is the symbol for it.

	In Fact H0 is True	**In Fact H0 is False**
Test Decides H0 True	Correct Decision	Type II Error
Test Decides H0 False	Type I Error	Correct Decision

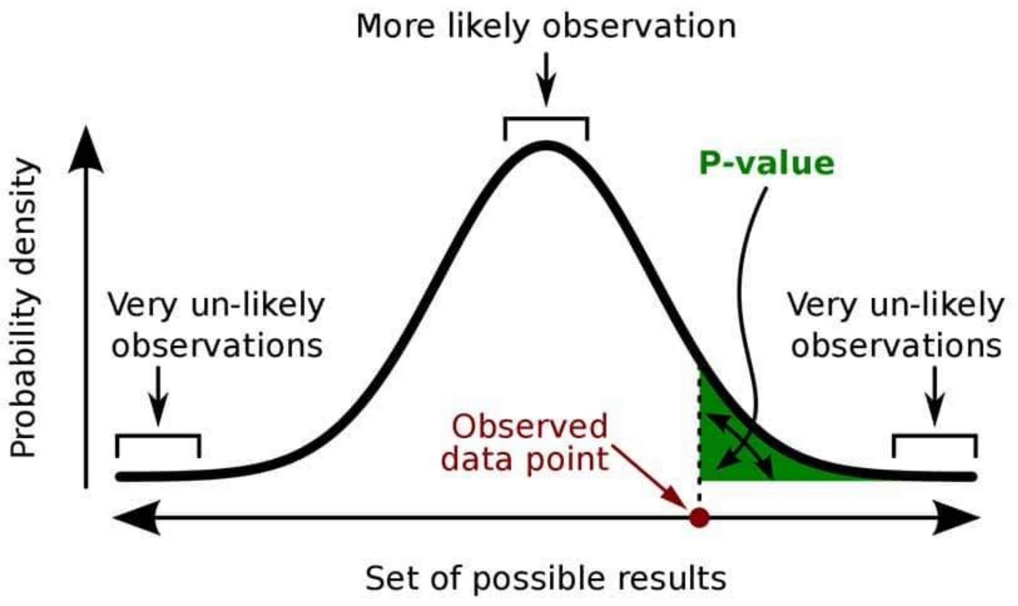
Decreasing the probability of a Type I error leads to an increase in the probability of a Type II error, and vice versa. A critical value (setting limits for a critical region) is usually chosen as a decision cut-off so that tests will have a small probability of Type I error. Errors depend on sample size.

https://www.simplilearn.com/tutorials/statistics-tutorial/hypothesis-testing-in-statistics

https://www.csus.edu/indiv/j/jgehrman/courses/stat50/hypthesistests/9hyptest.htm

Reminder – p-value

The **p-value**, or probability value, tells how likely it is that certain data could have occurred under the null hypothesis. **It is the probability of the observed outcome or something more extreme than the observed outcome, computed under the assumption that the null hypothesis, is true.** For instance, if your p-value is 0.05, that means that 5% of the time you would see a test statistic at least as extreme as the one you found if the null hypothesis was true.



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Note that p-values can be one-sided (figure) or two-sided if you count probabilities at both tails.

Statistical significance is another way of saying that the *p*-value of a statistical test is small enough to reject the null hypothesis of the test. The p-value can only tell whether or not the null hypothesis is supported (probabilistically) by measurements. It cannot tell whether the alternative hypothesis is true.

https://www.scribbr.com/statistics/p-value/
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6532382/

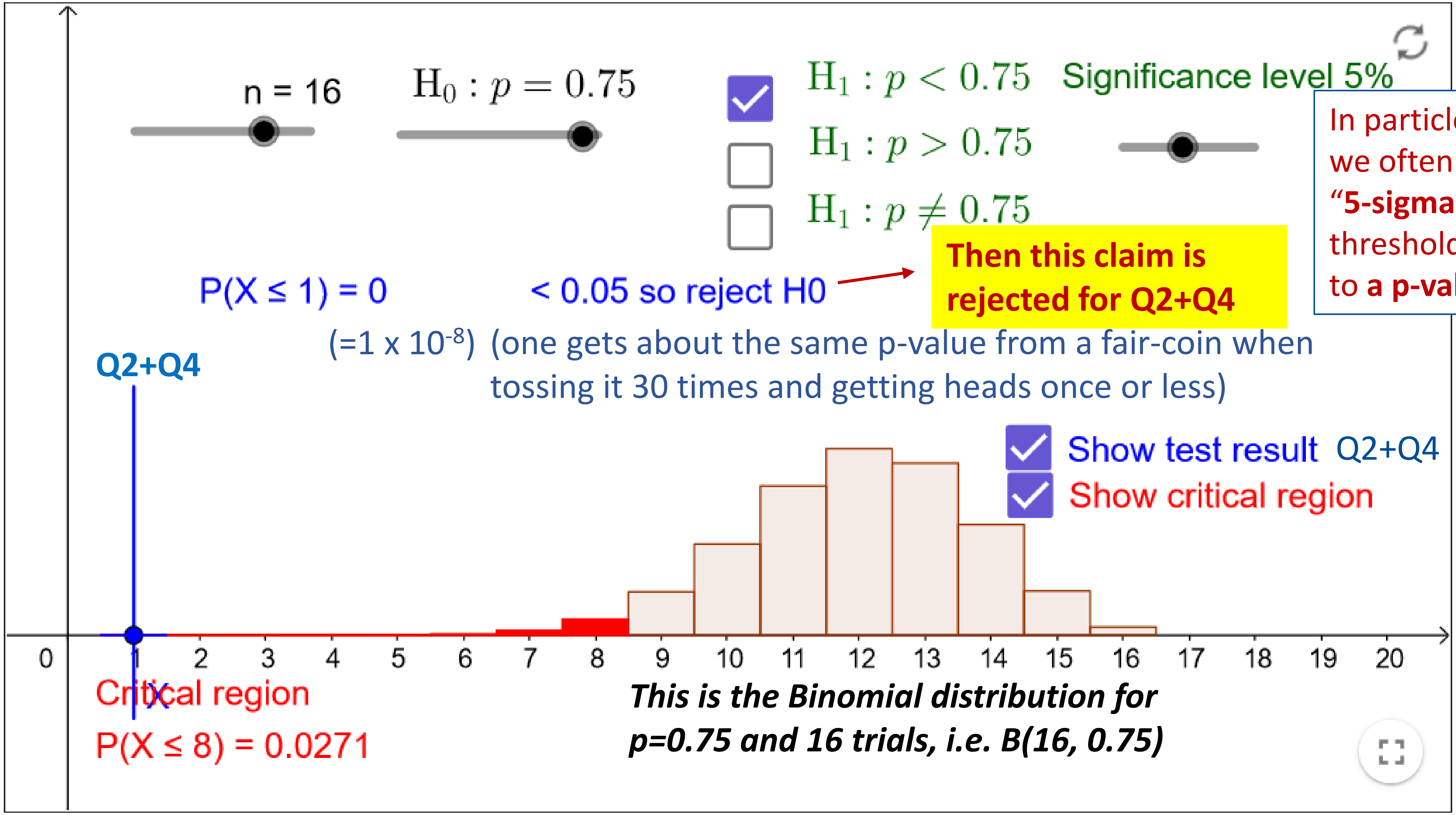
https://www.simplypsychology.org/p-value.html

Indian J Psychol Med. 2019 May-Jun; 41(3): 210–215.
doi: 10.4103/IJPSYM.IJPSYM_193_19

Testing H0 on Q2+Q4 data

Let's assume our Q1+Q3 measurements give the "true" probability of quenching – 0.75.

H0 : the probability of a coil to quench at least once below target current, with current gain below 600 A, is p = 0.75



Then this claim is rejected for Q2+Q4

($=1 \times 10^{-8}$) (one gets about the same p-value from a fair-coin when tossing it 30 times and getting heads once or less)

This is the Binomial distribution for p=0.75 and 16 trials, i.e. B(16, 0.75)

In particle physics <u>however</u> we often talk about "**5-sigma**" as a discovery threshold. This corresponds to **a p-value of (2 x) 3 x 10^{-7}**

https://home.cern/resources/faqs/five-sigma

https://www.geogebra.org/m/hcbw695z#material/x3B9QDSu

Testing "contingency"/dependency of two statistical samples

We don't know what the "true" probability **p** actually is and if such a number is meaningful at all. **Thus, the previous test is not too scientific**. But it depicts a common problem with a common solution and can be solved statistically.

In Statistics, a **contingency table** provides a way of portraying data that can facilitate calculating probabilities. The table displays sample values in relation to two different variables that may be dependent or contingent on one another.

For small samples, and especially for **2 x 2 tables**, there are two well-known "exact tests" for determining the equality of two (binomial) probabilities: Fisher's exact test and Barnard's exact test (both use assumptions).

One can also use a Log-Likelihood ratio test with Yates correction (for low statistics)

All tests use assumptions; the least biased among them, for 2x2 tables and low statistics, I believe is the **Barnard's test.**

The results are still the p-values we discussed. p-value limits ("significance level") are often 5% or 1% in medical studies; in sciences those levels are much lower. In Physics we often talk about **"evidence"** and **"observation"/"discovery"** limits – with p-value thresholds of **(2 x) 1.4 x 10^{-3}** , and **(2 x) 3 x 10^{-7}, respectively.** "(2 x)" comes from one- or two-sided choice.

In (particle) physics we often talk about "**5-sigma**" as a discovery threshold.
This corresponds to **a p-value of (2 x) 3 x 10^{-7}**.
(6σ is p-value of 2 x 1 x 10^{-9} / 4σ is p-value of 2 x 3.2 x 10^{-5} / 3σ is p-value of 2 x 1.4 x 10^{-3})

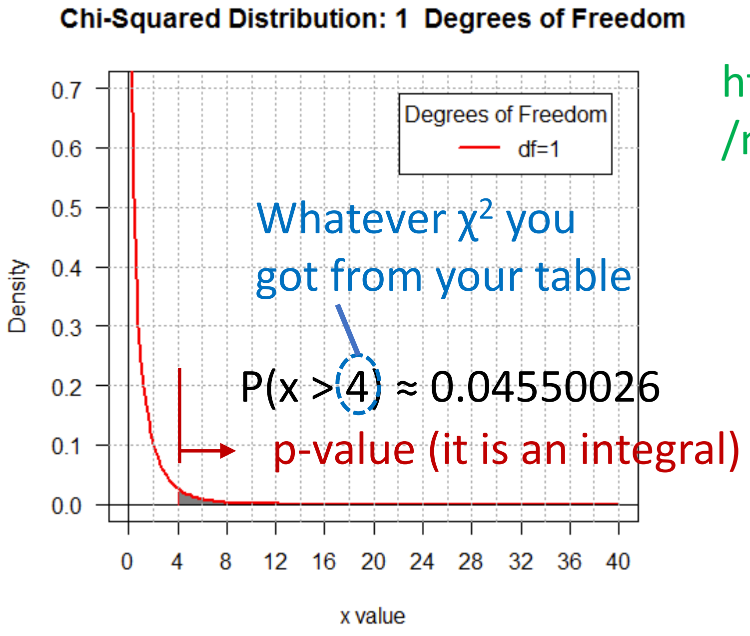
The "sigma" comes from Normal distribution approximation which is often a good approximation (at high statistics).

P-value from a contingency table – general logic

The p-value stemming from a single model distribution, like the Binomial test, seems straightforward. But what is we are comparing two data sets each with its own (presumably Binomial) distribution? One needs to construct test statistic, a good example is the χ^2 statistics, used here for tests of independence: $\chi^2_{df} = \sum \frac{(O_i - E_i)^2}{E_i}$ (**O**bserved and **E**xpected values, df is degrees of freedom).

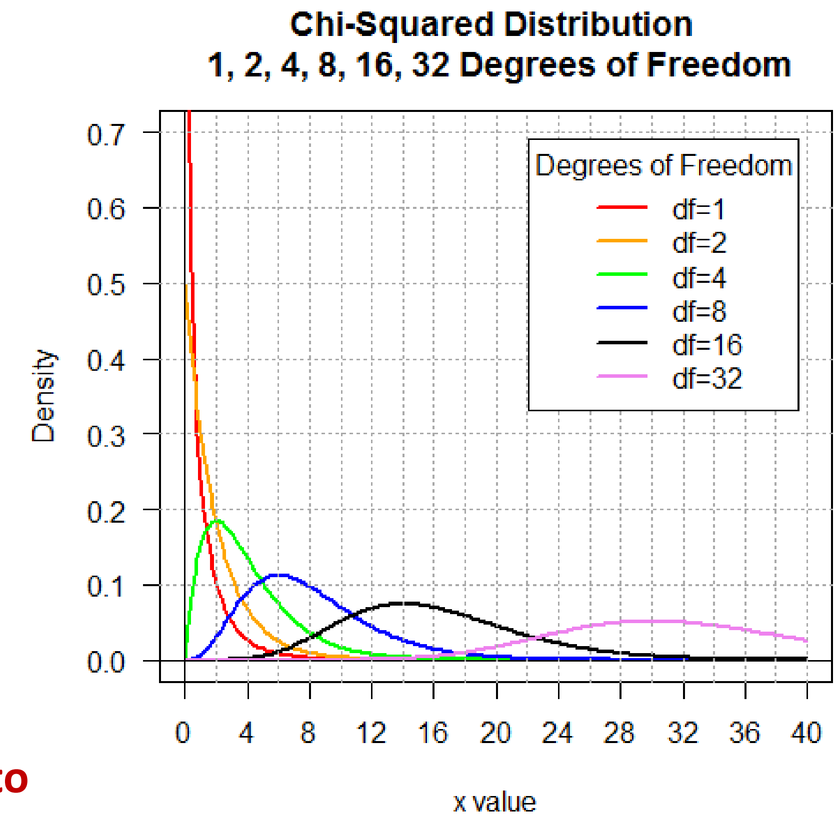
For a contingency table E is (raw total * column total) / total, generally different for each raw and column combination.

The df = (Nr -1)(Nc-1) for a table with number of rows Nr and number of columns Nc; thus, for 2x2 table **df = 1**.



https://courses.wccnet.edu/~palay/math160r/prob_chisq.htm

Not all test statistics are equally good given conditions. The bottom line is that the logic to get to the p-value is the same.



Back up: Ch2 from contingency tables

$$\chi^2_{df} = \sum \frac{(O_i - E_i)^2}{E_i}$$

O: observed

E: expected = (raw total x column total) (table total)
because:

Row total

	a	c	a+c
	b	d	b+d
Column total	a+b	c+d	N = a+b+c+d

(a+c)/N is a probability based on data (columns), (a+b)/N is a probability too (rows).
If rows and columns are independent, then the probability of one given the other is the multiplication of the two: (a+c) x (a+b)/ N^2 is also a probability. Multiplying this by the total number N gives the expected number based on that probability and the observed total:

E(1,1) = (a+c) x (a+b)/ N and so on for E(1,2), E(2,1), E(2,2)

Moreover, you can see that the Odds Ratio from the table with expected values (E11/E12)/(E21/E22) = 1, i.e. by construction chi2 takes the Null hypothesis to be correct. As any other test statistics.

Tests to use

Barnard's test

It examines all contingency tables with given number of events of the type considered (2x2 here).
It calculates probabilities using constraints from some test statistics ("discrepancy measure"), like Score or Wald, which depends on the true probability(s) in initial distributions – nuisance parameter, call it π.
Then it finds the maximum of p-value (π).

https://www.nbi.dk/~petersen/Teaching/Stat2009/Barnard_ExactTest_TwoBinomials.pdf

Mehta, Cyrus & Senchaudhuri, Pralay. (2003). Conditional versus Unconditional Exact Tests for Comparing Two Binomials.

Fisher's test

It considers all contingency tables of the type (2x2 here), which have the exact totals (sums of columns/rows) observed. Then the probability distribution of a table, among all possible tables, is the hypergeometric probability (given H0 is true) and it no longer depends on the nuisance parameter π. Thus, probabilities can be directly calculated, and p-value obtained.

G-test

Unlike the tests on the left, which calculate "exact" p-values, given conditions, it gives an approximate estimate of the p-value. It can be additionally corrected (like Yates correction). It is based on the log of the ratio of two likelihoods as the test statistic, which is χ^2-distributed; the p-value is estimated from integration over that distribution. Typically, it is not recommended for "small" numbers data.

https://search.r-project.org/CRAN/refmans/AMR/html/g.test.html

https://stats.libretexts.org/Bookshelves/Applied_Statistics/Biological_Statistics_(McDonald)/02%3A_Tests_for_Nominal_Variables/2.08%3A_Small_Numbers_in_Chi-Square_and_GTests

A bit more details on tests (much more details in refs)

Barnard's test

For a 2x2 contingency table, such as $X = [n_1, n_2; n_3, n_4]$, the normalized difference in proportions between the two categories, given in each column, can be written with pooled variance (Score statistic) as

$$T(X) = \frac{\hat{p}_2 - \hat{p}_1}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{c_1} + \frac{1}{c_2})}},$$

where $\hat{p} = (n_1 + n_3)/(n_1 + n_2 + n_3 + n_4)$, $\hat{p}_2 = n_2/(n_2 + n_4)$, $\hat{p}_1 = n_1/(n_1 + n_3)$, $c_1 = n_1 + n_3$ and $c_2 = n_2 + n_4$.
Alternatively, with unpooled variance (Wald statistic), the difference in proportions can we written as

$$T(X) = \frac{\hat{p}_2 - \hat{p}_1}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{c_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{c_2}}}.$$

The probability of observing X is

$$P(X) = \frac{c_1!c_2!}{n_1!n_2!n_3!n_4!} p^{n_1+n_2} (1-p)^{n_3+n_4},$$

where p is the unknown nuisance parameter.

Barnard's test considers all tables with category sizes c_1 and c_2 for a given p. The p-value is the sum of probabilities of the tables having a score in the rejection region, e.g. having significantly large difference in proportions for a two-sided test. The p-value of the test is the maximum p-value calculated over all p between 0 and 1.

https://search.r-project.org/CRAN/refmans/DescTools/html/BarnardTest.html
https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.barnard_exact.html

The p-value is then computed as:

$$\sum \binom{c_1}{x_{11}} \binom{c_2}{x_{12}} \pi^{x_{11}+x_{12}} (1-\pi)^{t-x_{11}-x_{12}}$$

where the sum is over all 2x2 contingency tables X such that: * $T(X) \leq T(X_0)$ when *alternative* = "less", * $T(X) \geq T(X_0)$ when *alternative* = "greater", or * $T(X) \geq |T(X_0)|$ when *alternative* = "two-sided". Above, c_1, c_2 are the sum of the columns 1 and 2, and t the total (sum of the 4 sample's element).

The returned p-value is the maximum p-value taken over the nuisance parameter π, where $0 \leq \pi \leq 1$.

Fisher's test

https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/fisher.test

For 2×2 cases, p-values are obtained directly using the (central or non-central) hypergeometric distribution.

For 2×2 tables, the null of conditional independence is equivalent to the hypothesis that the odds ratio equals one. 'Exact' inference can be based on observing that in general, given all marginal totals fixed, the first element of the contingency table has a non-central hypergeometric distribution with non-centrality parameter given by the odds ratio (Fisher, 1935). The alternative for a one-sided test is based on the odds ratio, so `alternative = "greater"` is a test of the odds ratio being bigger than `or`.

Two-sided tests are based on the probabilities of the tables, and take as 'more extreme' all tables with probabilities less than or equal to that of the observed table, the p-value being the sum of such probabilities.

Barnard's vs Fisher's (2x2 tables)

Barnard's test

(see back up)

Fisher's test

In most cases it will give a better estimate

More conservative and more popular (less computationally intensive)

It examines all contingency tables with given number of events of the type considered (2x2 here). It calculates probabilities using constraints from some test statistics ("discrepancy measure"), like Score or Wald, which depends on the true probability(s) in initial distributions – nuisance parameter, call it π. Then it finds the maximum of p-value (π).

It considers all contingency tables of the type (2x2 here), which have the exact totals (sums of columns/rows) observed. Then the probability distribution of a table, among all possible tables, is the hypergeometric probability (given H0 is true) and it no longer depends on the nuisance parameter π. Thus, probabilities can be directly calculated, and p-value obtained.

Getting COVID shots in same arm might result in stronger immune response

WBZ NEWS

BY MALLIKA MARSHALL, MD

AUGUST 23, 2023 / 5:50 PM / CBS BOSTON

eBioMedicine

Part of THE LANCET *Discovery Science*

Differences in SARS-CoV-2 specific humoral and cellular immune responses after contralateral and ipsilateral COVID-19 vaccination

Laura Ziegler • Verena Klemis • Tina Schmidt • Sophie Schneitler • Christina Baum • Jürgen Neumann • et al.

Show all authors • Show footnotes

Open Access • Published: August 11, 2023 • DOI: https://doi.org/10.1016/j.ebiom.2023.104743

analysed using a t-test (two-sided). Categorical analyses on vaccine responses were performed using the Fisher's test sex/gender and

spike-specific CD8 T-cell levels were significantly lower in the contralateral (0.026% (IQR 0.079%)) than in the ipsilateral group (0.063% (IQR 0.147), p = 0.004). As a result, the percentage of individuals with spike-specific CD8 T-cell levels above the detection limit was significantly lower after contralateral than after ipsilateral vaccination (34/79 (43.0%) versus 43/64 (67.2%), p = 0.004). CD4 and CD8 T-cell levels after polyclonal stimulation with *Staphylococcus aureus* Enterotoxin B (SEB) did not differ between the two groups (Fig. 1c), which indicates that the effect is vaccine-specific. In both

p-value

Software and reading material

(https://rdrr.io/cran/Barnard/man/barnard.test.html)

https://scistatcalc.blogspot.com/2013/11/barnards-test-calculator.html

https://www.r-statistics.com/2010/02/barnards-exact-test-a-powerful-alternative-for-fishers-exact-test-implemented-in-r/

https://search.r-project.org/CRAN/refmans/DescTools/html/BarnardTest.html
https://rpubs.com/juanhklopper/g_tests_for_categorical_variables
https://influentialpoints.com/Training/g-likelihood_ratio_test.htm
https://www.statology.org/fishers-exact-test/

https://rdrr.io/cran/DescTools/

https://rdrr.io/snippets/

"R" Project for Statistical Computing
and some on-line tools were used
(and cross-checked with known calculations)

You can execute the code on the right in your browser.

```
library(DescTools)
tab <- as.table(matrix(c(2, 2, 7, 5), nrow=2,
          dimnames=list(treat=c("I","II"), out=c("I","II"))))

ft <- fisher.test(tab, alternative = "l")
(bt <-BarnardTest(tab, method = "z-pooled", alternative = "l", fixed = "NA"))

cat("\nFisher's test (one-sided): ", ft$p.value)
cat("\nBarnard's test (one-sided): ", bt$p.value)
```

Fisher's
Barnard's
Log-likelihood

```
ft2 <- fisher.test(tab, alternative = "two.sided")
bt2 <-BarnardTest(tab, method = "z-pooled", alternative = "two.sided")
llt <- GTest(tab, correct = "yates")

cat("\nFisher's test (two-sided): ", ft2$p.value)
cat("\nBarnard's test (two-sided): ", bt2$p.value)
cat("\n Log-likelihood ratio test /a.k.a. G-test/ with Yates' correction: ", llt$p.value)
```

(a lot of refs and examples exist for Fisher's test)

Back up: More ref

John Ludbrook, Analysis of 2 × 2 tables of frequencies: matching test to experimental design, *International Journal of Epidemiology*, Volume 37, Issue 6, December 2008, Pages 1430–1435, https://doi.org/10.1093/ije/dyn162

https://www.nbi.dk/~petersen/Teaching/Stat2009/Barnard_ExactTest_TwoBinomials.pdf

Conditional versus Unconditional Exact Tests for Comparing Two Binomials

Cyrus R. Mehta and Pralay Senchaudhuri

4 September 2003

https://scistatcalc.blogspot.com/2013/11/barnards-test-calculator.html

Eur J Epidemiol (2016) 31:337–350
DOI 10.1007/s10654-016-0149-3

CrossMark

ESSAY

Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations

Sander Greenland[1] · Stephen J. Senn[2] · Kenneth J. Rothman[3] · John B. Carlin[4] · Charles Poole[5] · Steven N. Goodman[6] · Douglas G. Altman[7]

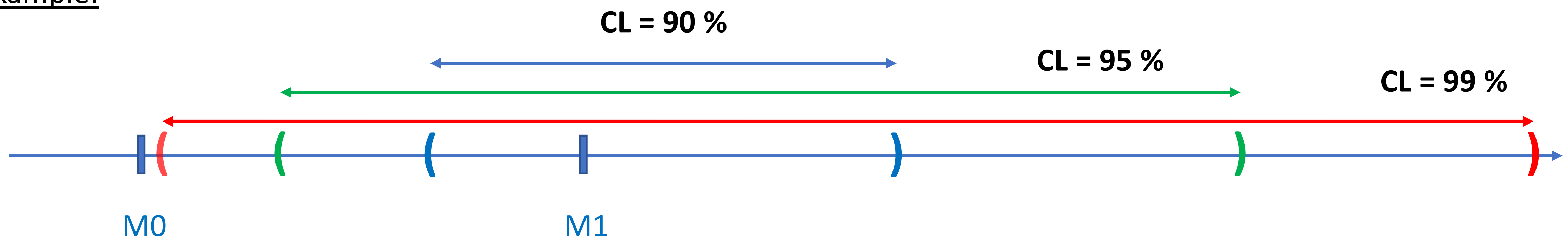
Chiba Y (2016) A Note on Exact Tests and Confidence Intervals for Two-by-Two Contingency Tables in Randomized Trials. Int J Clin Res Trials 1: 102. doi: https://doi.org/10.15344/2456-8007/2016/102

41

A word about Confidence Intervals

Let’s M1 is the mean of a measurement or just a measurement with some statistical uncertainty.
Is M1 statistically consistent with the M0 value?

Example:



If M0 is the one from the Null hypothesis, then the **Confidence Interval (CI)** is the interval where the probability to find the true M is equal to CL (confidence level) – based on observations and test statistics. CI relates to how far away the Null Hypothesis is from observations in terms of uncertainties (CI ~ observation ± normalized [CL x uncertainty]).

CI provide(s) more information than p-values alone, the interval(s) is(are) given in the units of the measurements.
It is often recommended to quote both CI and obtained p-value.

Too low statistics for some tests

For some tests, data are simply not enough to answer with good precision.
For instance, this table we saw

	C2 in Q1/Q3	C2 not in Q1/Q3
Not quenched	**5**	**2**
Quenched	**9**	**0**

gives this answer

```
        Fisher's Exact Test for Count Data

data:  tab
p-value = 0.175
alternative hypothesis: true odds ratio is less than 1
95 percent confidence interval:
 0.000000 2.567235
sample estimates:
odds ratio
         0
```

If we double the data in each box (just to give an example):

	C2 in Q1/Q3 (example)	C2 not in Q1/Q3 (example)
Not quenched	**10**	**4**
Quenched	**18**	**0**

then we'll get

```
        Fisher's Exact Test for Count Data

data:  tab
p-value = 0.02784
alternative hypothesis: true odds ratio is less than 1
95 percent confidence interval:
 0.0000000 0.7592865
sample estimates:
odds ratio
         0
```

Remember that H0 meant the odds ratio is consistent with 1.
Wide confidence intervals (like containing the whole [0,1] interval) suggest low statistical power.

C1 = FNAL coil
C2 = BNL coil

Confidence intervals – a note and an example

	Q1+Q3	Q2+Q4
Not quenched	**5**	**13**
Quenched	**13**	**3**
Fraction quenched	0.72	0.188
Odds	5/13=0.385	13/3 = 4.333

	P-value for one-sided test	P-value for two-sided test
Fisher's test	**2.4×10^{-3}**	**2.6×10^{-3}**
Barnard's test	**9.9×10^{-4}**	**1.9×10^{-3}**
G-test(Yates' corr.)		**4.6×10^{-3}**

There are statistical/computational reasons to also use the Odds Ratio (**OR**) which, of course, is the ratio of **odds**: **OR** = 0.385/4.333 = 0.089, in our case. Statistical tests also give confidence intervals around the **computed OR** (which is corrected based on the statistics used). Confidence interval, typically 95%, means that based on the test statistics and data, we'll get the **true OR** within this interval 95% of the times. Note that **OR** = 1 implies fractions between groups are the same, i.e. the usual H0 is true. If we are to discover categories which are related to each other (low p-value) this also means the confidence intervals of **OR** will not contain 1. However, how close we are to 1 and how wide the confidence interval is, gives more info than just the p-value.

```
        Fisher's Exact Test for Count Data

data:  tab
p-value = 0.00235
alternative hypothesis: true odds ratio is less than 1
95 percent confidence interval:
 0.0000000 0.4445466
sample estimates:
odds ratio
0.09692987
```

The interval is not symmetric

It differs a bit from simple calculations, it is based on data + test statistics

Statistical tests

HYPOTHESIS TESTING FOR QUENCHING OF COILS IN Q1||Q3 VS Q2||Q4

Conditions and Data	Test	PV (1-S)	PV(2-s)	95% CI
("non-weak", "<300 A"); 6/12 vs 16/0	FT BT	3.4×10^{-5} 2.1×10^{-5}	3.7×10^{-5} 4.1×10^{-5}	(0.0,0.15)
("non-weak", "<600 A"); 5/13 vs 16/0	FT BT	9.2×10^{-6} 3.6×10^{-6}	9.8×10^{-6} 7.2×10^{-6}	(0.0,0.12)
("non-weak", "no limit"; 5/13 vs 13/3	FT BT	2.4×10^{-3} 9.9×10^{-4}	2.6×10^{-3} 1.9×10^{-3}	(0.0,0.44)
("all", "<300 A"); 6/13 vs 16/2	FT BT	4.7×10^{-4} 1.7×10^{-4}	6.3×10^{-4} 2.8×10^{-4}	(0.0,0.32)
("all", "<600 A"); 5/14 vs 16/2	FT BT	1.4×10^{-4} 4.8×10^{-5}	1.9×10^{-4} 8.2×10^{-5}	(0.0,0.26)
("all", "no limit"); 5/14 vs 13/5	FT BT	6.4×10^{-3} 3.9×10^{-3}	8.6×10^{-3} 7.7×10^{-3}	(0.0,0.58)
("non-weak", "<600 A" + first quenches) ; 5/13 vs 10/6	FT BT	4.50×10^{-2} 3.09×10^{-2}	8.6×10^{-3} 7.7×10^{-3}	(0.0,0.97)

For each data condition, we show the 2x2 contingency table in the notations introduced earlier in the text. FT and BT (tests) are performed giving 1-s and 2-s PV and 95% CI (FT, 1-s based). Non-rejected 1-s H0's are highlighted.

Article DOI: 10.1109/TASC.2023.3341871
Manuscript Number: MT28-4PoM03-09

Dividing data in sub-categories gives us more insights but statistical power is reduced. We investigate four more H0's, all for the tier (**"non-weak", "< 600 A"**) : (I) a first quenched coil in a magnet has other training quenches independently on where it is in the magnet (Q1||Q3 vs Q2||Q4); (II) for quenches in Q1||Q3 only: there is no relation between quenching and coil origin (FNAL coils vs BNL coils); (III) for BNL coils only: there is no relation between quenching and where coils are placed in a magnet (Q1||Q3 vs Q2||Q4); (IV) for FNAL coils only: there is no relation between quenching and where coils are placed in a magnet (Q1||Q3 vs Q2||Q4). Results are in Table III.

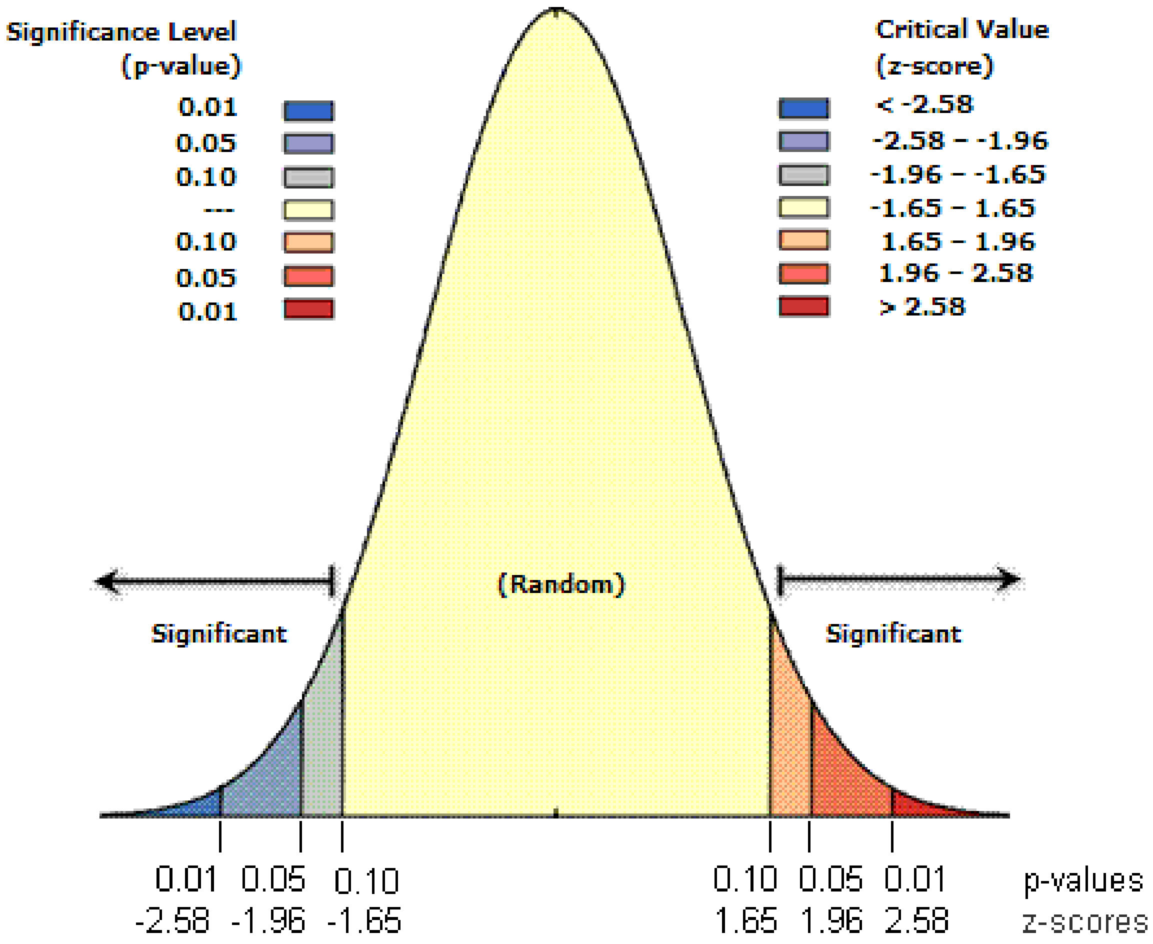
HYPOTHESIS TESTING (SUBSETS)

Hypothesis and data	Test	PV (1-s)	PV(2-s)	95% CI
(I) ; 0/2 vs 6/0	FT BT	3.6×10^{-2} 1.1×10^{-2}	3.6×10^{-2} 1.1×10^{-2}	(0.0,0.79)
(II) ; 0/4 vs 5/9	FT BT	0.23 0.11	0.28 0.20	(0.0,2.83)
(III) ; 5/9 vs 2/0	FT BT	0.18 0.12	0.18 0.12	(0.0,2.56)
(IV) ; 0/4 vs 14/0	FT BT	3.3×10^{-4} 7.2×10^{-5}	3.3×10^{-4} 7.2×10^{-5}	(0.0,0.13)

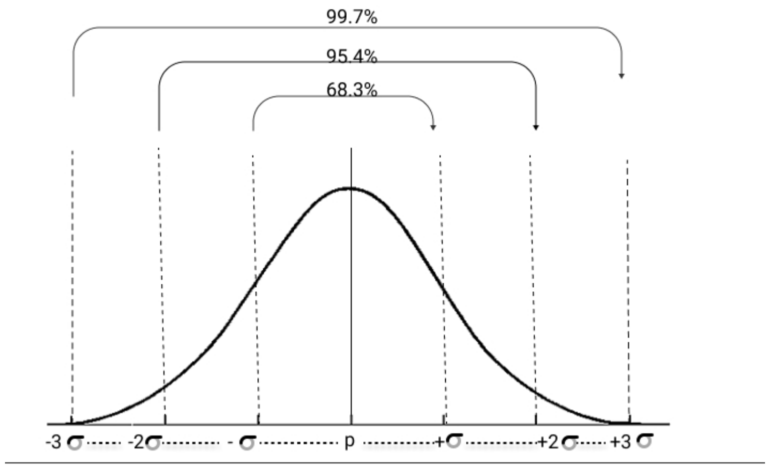
For each hypothesis, we show the 2x2 contingency table in the notations introduced earlier in the text. Testing and results follow notations in Table II.

Back up : P-values, CI, σ

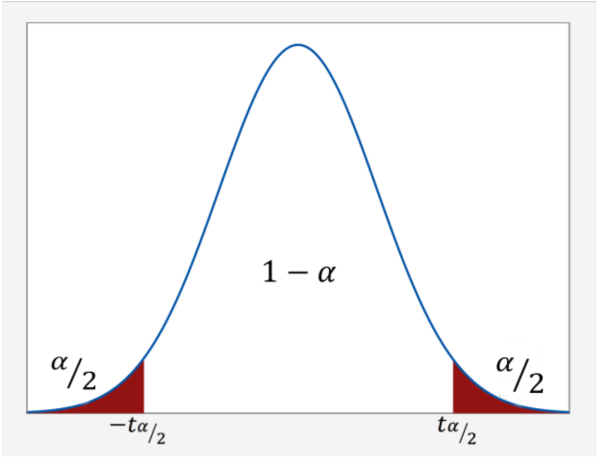
https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/what-is-a-z-score-what-is-a-p-value.htm



CI



https://www.analyticsvidhya.com/blog/2022/01/understanding-confidence-intervals-with-python/



https://online.stat.psu.edu/statprogram/reviews/statistical-concepts/confidence-intervals

P-values and CI (confidence intervals) are about integrals,
sigma is a function parameter making interpretation easier
(it is the standard deviation of the Normal distribution)

Be aware of two- or one-sidedness for integrals.