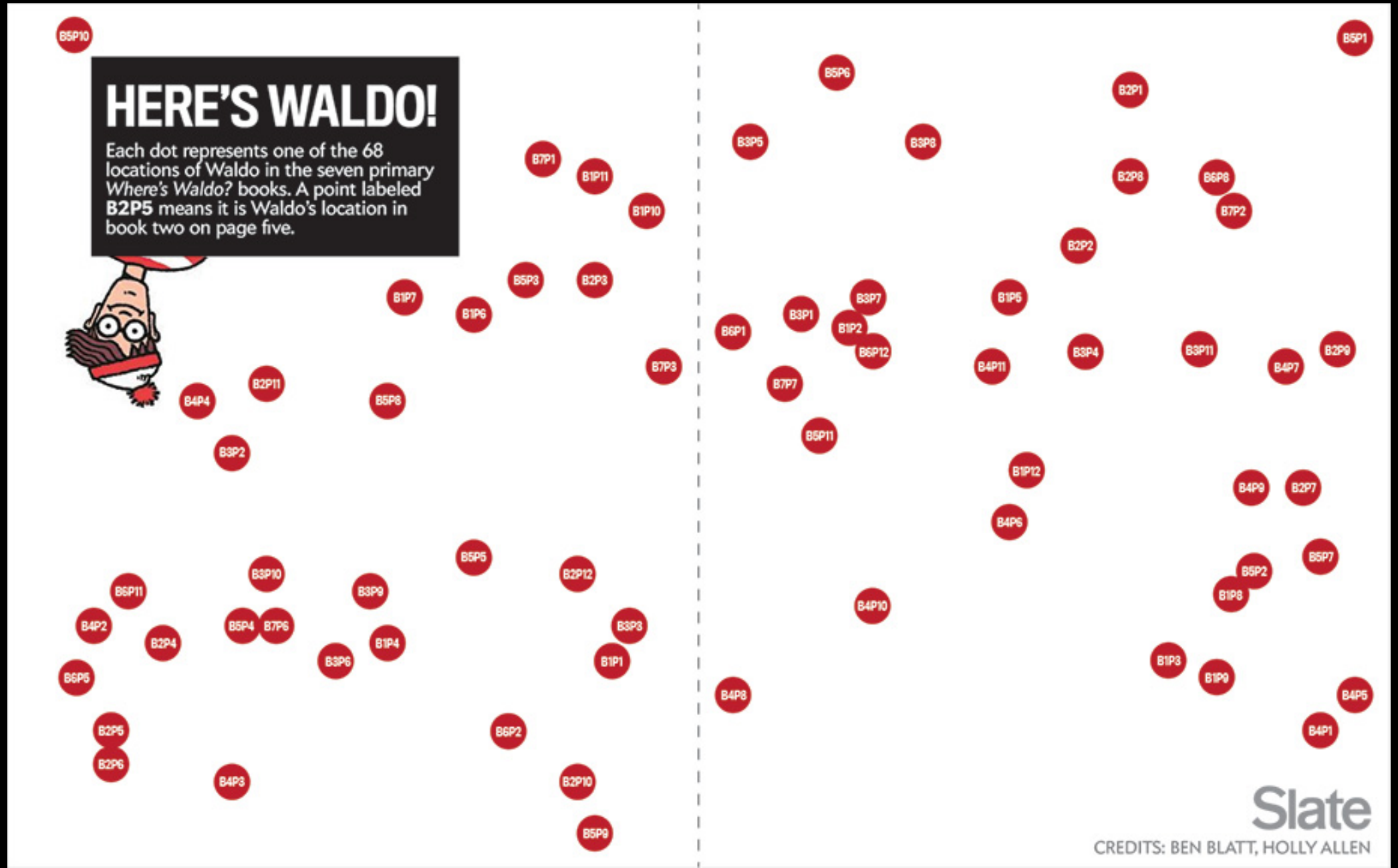
Hypothesis Testing II

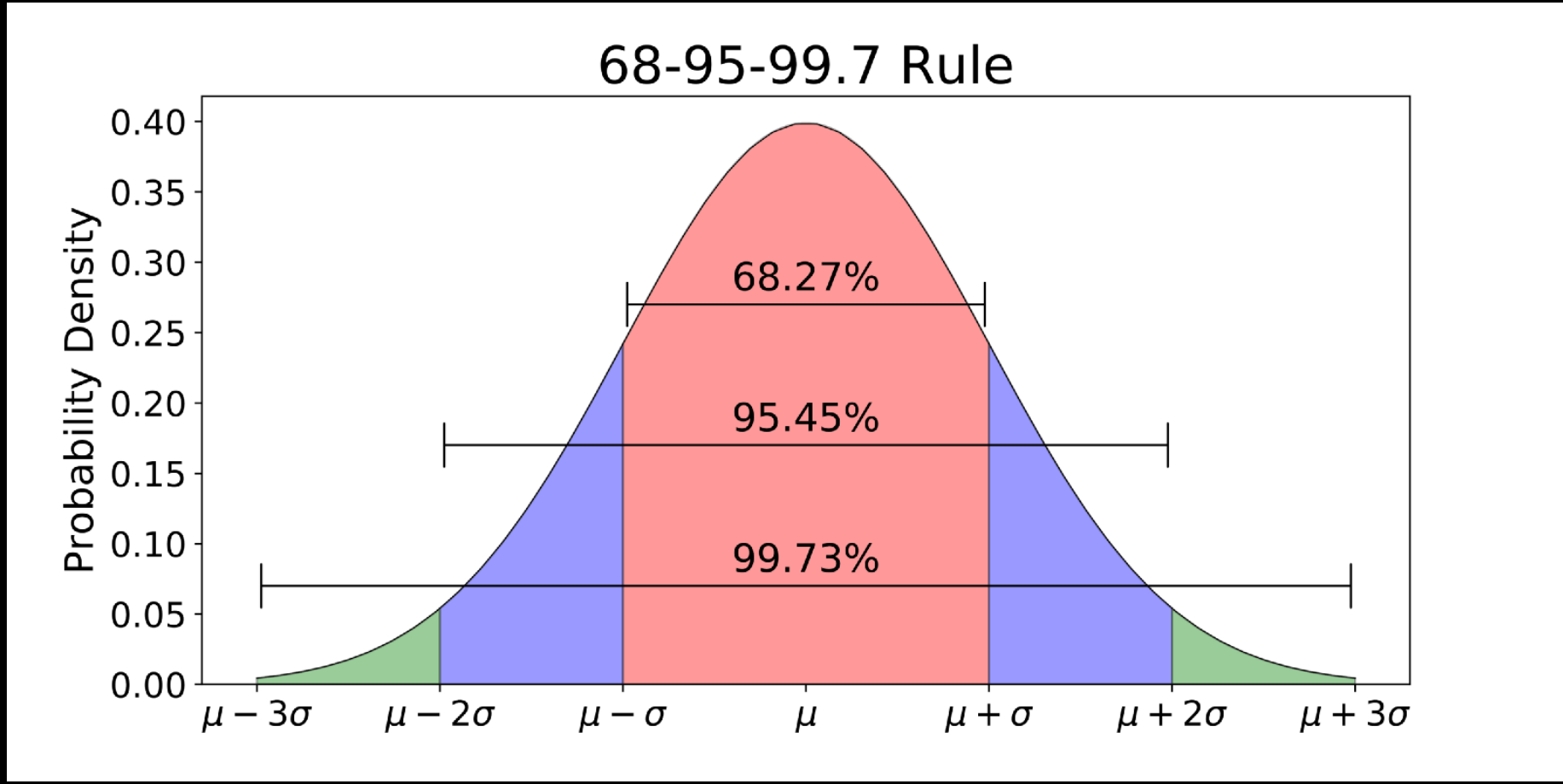
?10/10/1977?


DOES MICHAEL BAY HATE your childhood?
CONFIRMED
iDeathNoteBook Air
KISSING BOOTH FOR CHARITY
Game on Tables
FOOD
SELL UR USED GAMES HERE
USED GAMES $99.99
DISKINECT
Magic show!
Henching!
PLEASE COSPLAY US
this way
BITE OFF THEIR LIMBS
EXIT
WORKSHOP
HOW TO: Zombie COSPLAY
WORKSHOP
HOW TO: KILL A Zombie
PXL CON
COSPLAY COMPETITION
FREAKSHOW COLLECTIBLES
LIFE-LIKE FIGURINES!
DEAD
ZOMBIE-KILLING WORKSHOP AHEAD!!!
CAUTION! MINE
DO NOT CROSS THE STREAMS!
tase us bro
Pink Unicorn Comics
Buff
Vampire slayer
Maid Cafe
I ♥ YAOI
SONY
BLEACH
KILLS mold + germs + Evil Spirits
Closed Due to our booth being stolen :(
Wow!!
TRY OUT REAL-LIFE GADGETS FROM MOVIES GAMES & TV!
GADGET ZONE
Petting Zoo
First Aid
PXL CON

Waldo Stats. Is this a normal distribution?



Standard deviations and probability of the population mean



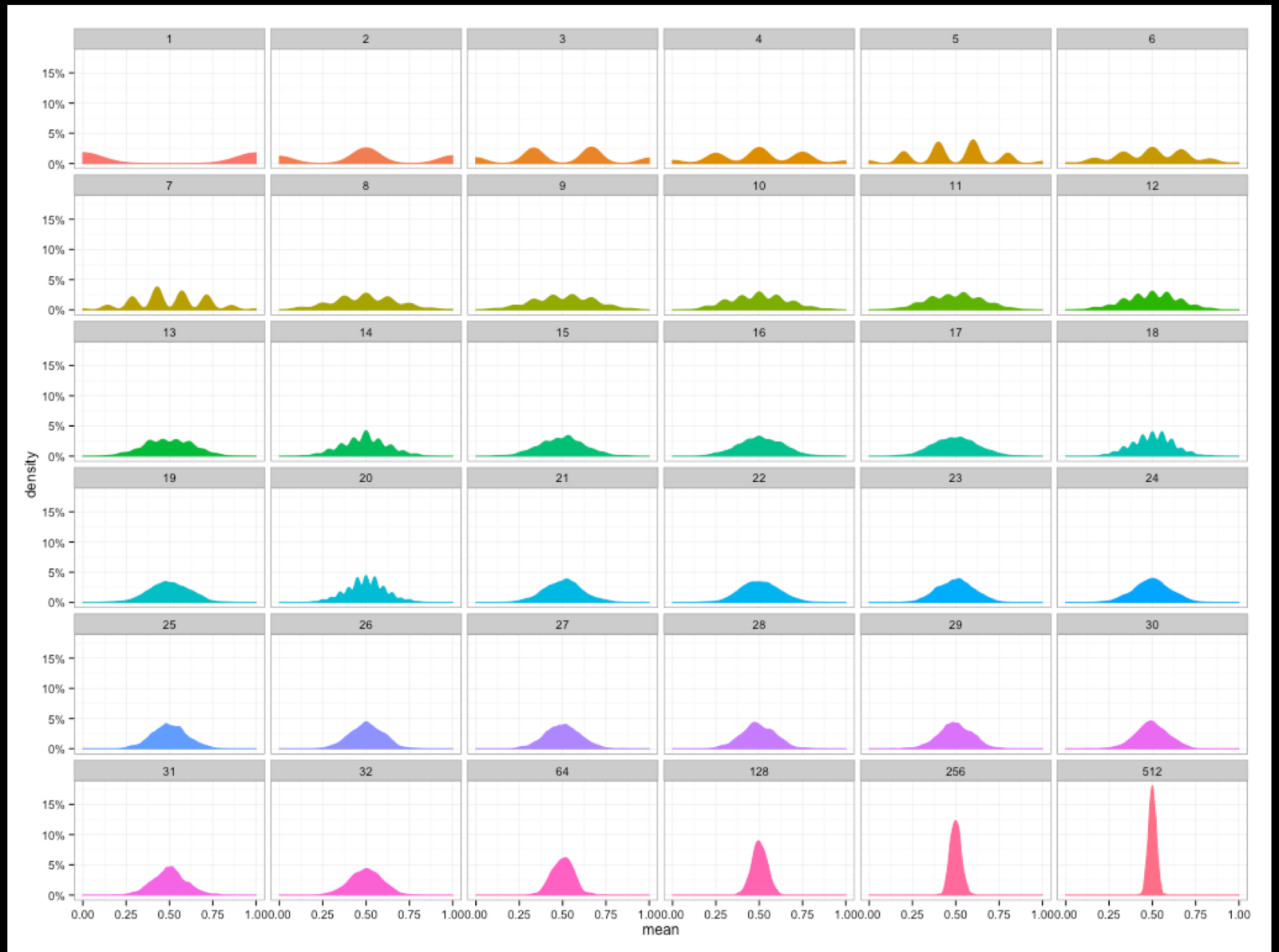
The process of testing hypothesis

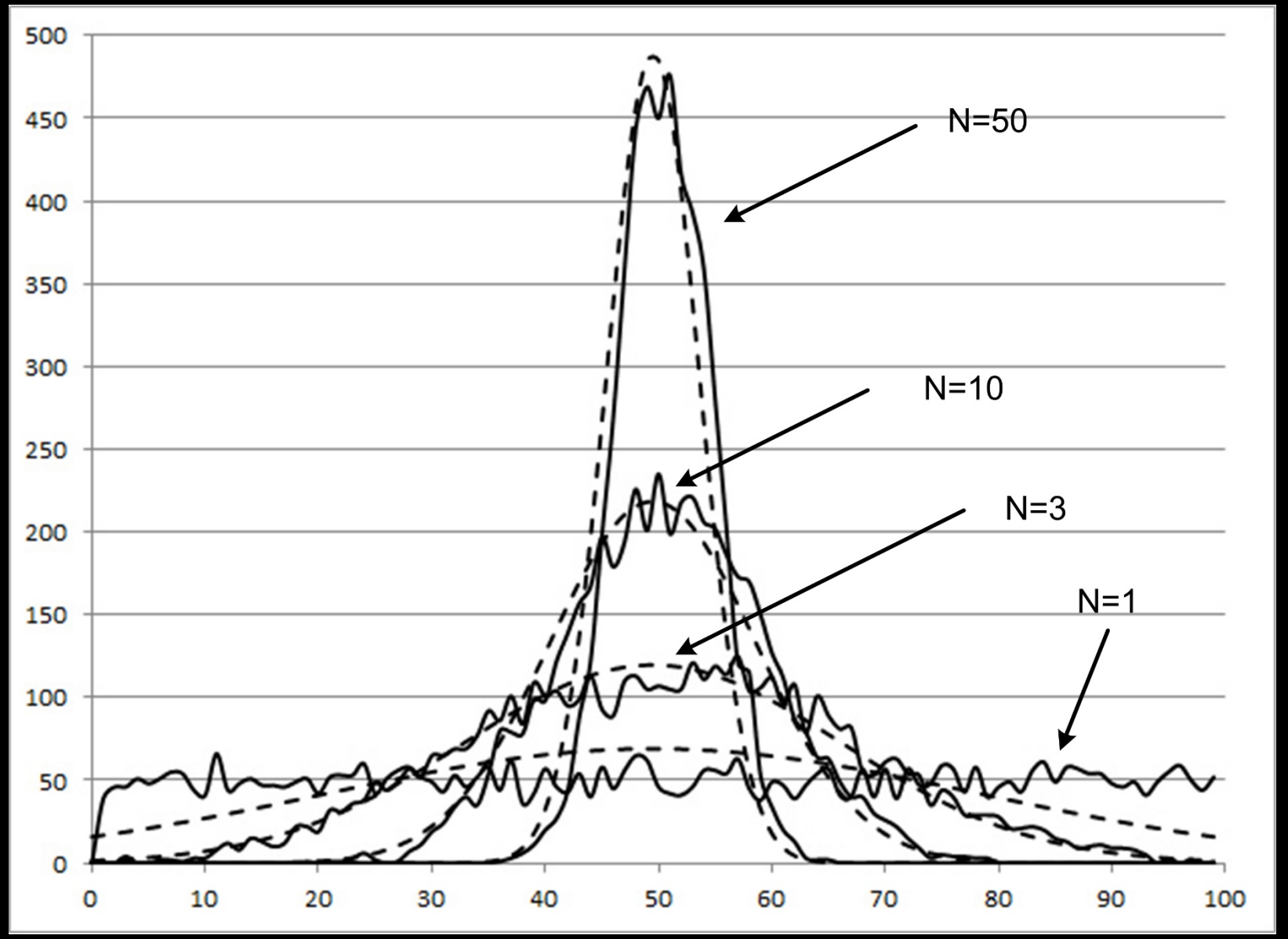
1. Examine null hypothesis
2. If suspicious, try to disprove it (where's the confidence numbers)
3. Decide how to sample randomly
4. Figure out how many N to sample?
5. Calculate mean
6. Calculate standard deviation
7. Calculate Z score (how many deviations away are you)
8. Reject or accept based on probability usually >=95%, definitely if over 99.7%

How many N to sample?

More is better

1. Cost + time a factor
2. As N increases the standard deviation decreases
3. Confidence interval decreases, +- error




500
450
400
350
300
250
200
150
100
50
0
0
10
20
30
40
50
60
70
80
90
100
N=50
N=10
N=3
N=1

Z-Values

A z-value gives you p-value (probability of getting this value) This is a cumulative from the mean so multiply your Z*2 to get a total probability. E.g., z=2 → p = .47*2 = 95%

z	+0.00	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0.2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0.3	0.11791	0.12172	0.12552	0.12930	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0.4	0.15542	0.15910	0.16276	0.16640	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0.5	0.19146	0.19497	0.19847	0.20194	0.20540	0.20884	0.21226	0.21566	0.21904	0.22240
0.6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.25490
0.7	0.25804	0.26115	0.26424	0.26730	0.27035	0.27337	0.27637	0.27935	0.28230	0.28524
0.8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891
1.0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1.1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900	0.38100	0.38298
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1.3	0.40320	0.40490	0.40658	0.40824	0.40988	0.41149	0.41308	0.41466	0.41621	0.41774
1.4	0.41924	0.42073	0.42220	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1.5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1.6	0.44520	0.44630	0.44738	0.44845	0.44950	0.45053	0.45154	0.45254	0.45352	0.45449
1.7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.46080	0.46164	0.46246	0.46327
1.8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1.9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558	0.47615	0.47670
2.0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077	0.48124	0.48169
2.1	0.48214	0.48257	0.48300	0.48341	0.48382	0.48422	0.48461	0.48500	0.48537	0.48574
2.2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2.8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861
3.0	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.49900
3.1	0.49903	0.49906	0.49910	0.49913	0.49916	0.49918	0.49921	0.49924	0.49926	0.49929
3.2	0.49931	0.49934	0.49936	0.49938	0.49940	0.49942	0.49944	0.49946	0.49948	0.49950
3.3	0.49952	0.49953	0.49955	0.49957	0.49958	0.49960	0.49961	0.49962	0.49964	0.49965
3.4	0.49966	0.49968	0.49969	0.49970	0.49971	0.49972	0.49973	0.49974	0.49975	0.49976
3.5	0.49977	0.49978	0.49978	0.49979	0.49980	0.49981	0.49981	0.49982	0.49983	0.49983
3.6	0.49984	0.49985	0.49985	0.49986	0.49986	0.49987	0.49987	0.49988	0.49988	0.49989
3.7	0.49989	0.49990	0.49990	0.49990	0.49991	0.49991	0.49992	0.49992	0.49992	0.49992
3.8	0.49993	0.49993	0.49993	0.49994	0.49994	0.49994	0.49994	0.49995	0.49995	0.49995
3.9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997
4.0	0.49997	0.49997	0.49997	0.49997	0.49997	0.49997	0.49998	0.49998	0.49998	0.49998

T-Values

If N <30 then the standard deviations become larger since n is smaller. So it now takes 2.262 deviations to get 95% of the data.

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Any statistic has an error +- (we're speculating)

We call this the margin of error

For the mean we have:

Mean +-Z*(Standard Deviation/sqrt(N))

$$\bar{X} \pm Z\frac{s}{\sqrt{n}}$$

E.g., for the Query analysis we did we have

Mean = 4, Z=-3.40, S=1.76, so we have

4 +- 1.09 @99.97%. So we can say we are 999.7% confident that population mean is 4 +- 1.09 or we are 999.7% confident that the population mean is within the confidence interval 2.91 and 5.09

Margin of Error continued

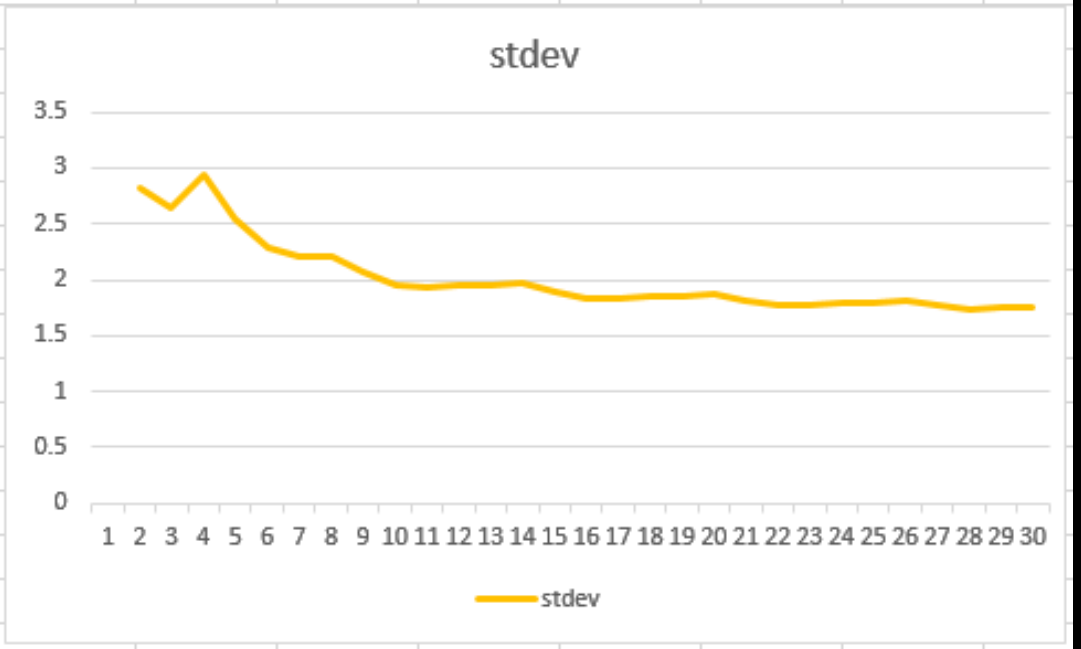
$$\bar{x} \pm z\frac{s}{\sqrt{n}}$$

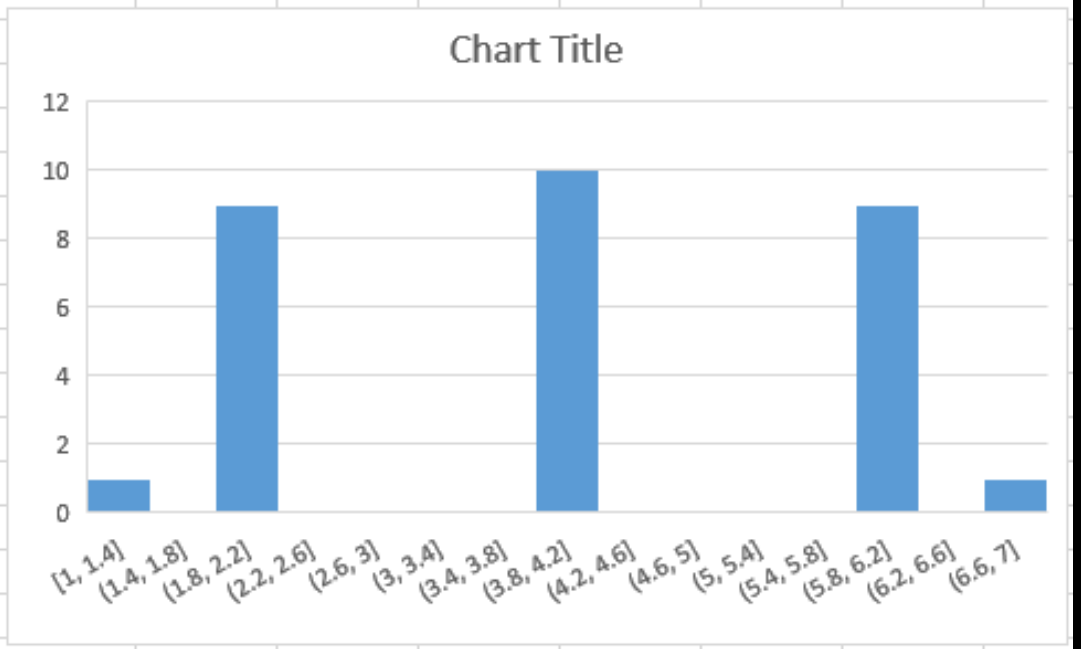
We could have chosen a 95% margin of error also and by doing so our confidence interval would have decreased. E.g.,

Mean = 4, Z=2, S=1.76, so we have

4 +- .64 @95%. So we can say we are 95% confident that population mean is 4 +- .64 or we are 95% confident that the population mean is within the confidence interval 3.36 and 4.64. This is how you should present your findings. Statistic, N, Standard Deviation and confidence numbers.

null hypothesis						
10	running mean	measure nearest minute	n	stdev	z-score	confidence
	2	2	1			0
	4	6	2	2.828427	-2.12132	-4.24264
	3.5	1	3	2.645751	-2.456769	-3.75278
	4	7	4	2.94392	-2.038099	-3
	5.5	4	5	2.54951	-1.765045	-2.01246
	4	4	6	2.280351	-2.631174	-2.44949
	5	6	7	2.21467	-2.257673	-1.88982
	4	2	8	2.203893	-2.722456	-2.12132
	3	4	9	2.061553	-3.395499	-2.33333
	4	4	10	1.943651	-3.086975	-1.89737
	3	2	11	1.940009	-3.60823	-2.11058
	4	6	12	1.954017	-3.070598	-1.73205
	6	6	13	1.951331	-2.049883	-1.1094
	4	2	14	1.961161	-3.059412	-1.60357
	3	4	15	1.889822	-3.704052	-1.80739
	4	4	16	1.825742	-3.286335	-1.5
	3	2	17	1.833111	-3.818646	-1.69775
	4	6	18	1.847096	-3.248342	-1.41421
	6	6	19	1.852768	-2.158932	-0.91766
	4	2	20	1.863782	-3.21926	-1.34164
	3	4	21	1.81659	-3.853373	-1.52753
	4	4	22	1.772811	-3.384456	-1.2792
	3	2	23	1.781548	-3.929167	-1.4596
	4	6	24	1.793709	-3.345025	-1.22474
	6	6	25	1.800926	-2.22108	-0.8
	4	2	26	1.811077	-3.312946	-1.1767
	3	4	27	1.775907	-3.941648	-1.34715
	4	4	28	1.74271	-3.442914	-1.13389
	3	2	29	1.751143	-3.997389	-1.29987
final	4	6	30	**1.761661**	**-3.405877**	**-1.09545**





Type I & Type II errors

Null Hypothesis True	Did we reject	Example	Outcome	Probability we're right	Probability wrong
t	t	They say 10 minutes is the best, we say 4 but in fact they're correct (maybe we needed more N)	Type 1 error	99.970%	**0.030%**
t	f	They say 10 minutes is the best, we say 9 but 9 is within 95% so the null hypothesis holds.	Great, should have accepted it	**94.900%**	5.100%
f	f	They say 10 minutes is the best, we say 9 but 9 is within 95% so the null hypothesis holds. We are both wrong proven with higher N or another study is conducted and finds we sampled poorly	Type 2 error	94.900%	**5.100%**
f	t	They say 10 minutes is the best, we say 4 and we're correct (we increased N later on and verified)	Great, should have rejected it	**99.970%**	0.030%

Some examples

- The average time it takes to find Waldo at Comic Con by a random group of people over age 6 is 7 minutes. We surveyed 30 random people over age six and they found Waldo in an average time of 5 minutes. Who is right?
- Null hypo = mean = 7
- New hypo = mean = 5. N = 30, calculated sdev = 2.08, Z-Score = -.96
- Z-table look up gives = .33*2 → p = 66%. So there's a 66% chance that the population mean is in this range. Margin of error = 5+-.36 66% confidence. Should we reject?

Time permitting in class examples with Waldo