Στατιστικός έλεγχος υποθέσεων (Μέρος 2ο)

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη διακύμανση **σ^2** ενός κανονικού πληθυσμού με ένα τυχαίο δείγμα μεγέθους **n**

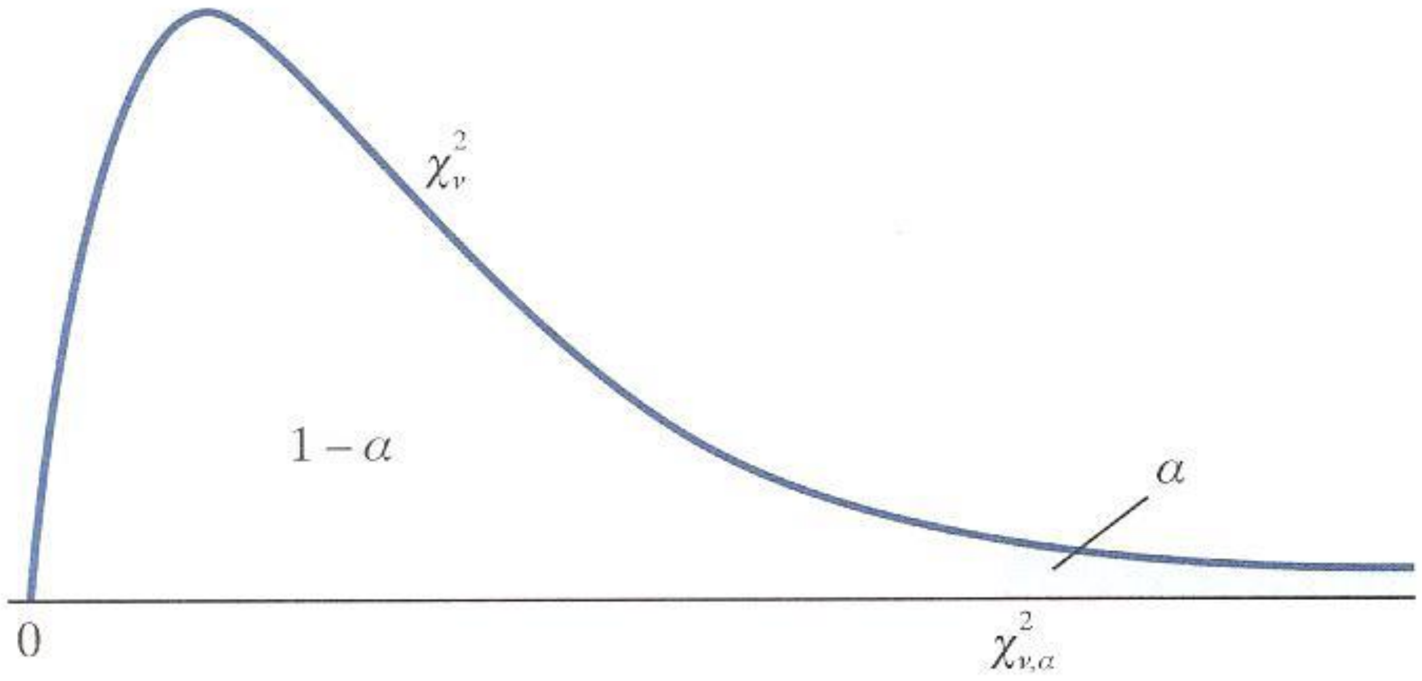
H_0: $\sigma^2 = \sigma_0^2$

Περιοχή απόρριψης της H_0

H_1: $\sigma^2 \neq \sigma_0^2$	H_1: $\sigma^2 > \sigma_0^2$	H_1: $\sigma^2 < \sigma_0^2$
$\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{n-1;\alpha/2}$ ή $\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{n-1;1-\alpha/2}$	$\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{n-1;a}$	$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{n-1;1-a}$

Η τ.μ. $\frac{(n-1)S^2}{\sigma_0^2}$, για κανονικό πληθυσμό, προσεγγίζεται ικανοποιητικά από την κατανομή χ^2 με n-1 βαθμούς ελευθερίας

Κατανομή χ^2 με *n* βαθμούς ελευθερίας – καμπύλη



$$P(X > \chi^2_{n;a}) = a$$

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη διαφορά **$\mu_1 - \mu_2$** των μέσων τιμών δύο πληθυσμών με δύο ανεξάρτητα δείγματα μεγέθους **n_1** και **n_2** αντίστοιχα

H_0: $\mu_1 - \mu_2 = \delta$

- Οι διακυμάνσεις **σ_1^2** και **σ_2^2** είναι **γνωστές** και οι πληθυσμοί είναι **κανονικοί**
- Οι διακυμάνσεις **σ_1^2** και **σ_2^2** είναι **γνωστές** και τα μεγέθη των δειγμάτων **n_1** και **n_2** είναι μεγάλα **(≥30)** (οτιδήποτε πληθυσμός) [*έλεγχος επιπέδου σημαντικότητας α κατά προσέγγιση*]

Περιοχή απόρριψης της H_0

H_1: $\mu_1 - \mu_2 \neq \delta$	H_1: $\mu_1 - \mu_2 > \delta$	H_1: $\mu_1 - \mu_2 < \delta$
$\lvert Z \rvert = \dfrac{\lvert \bar{X} - \bar{Y} - \delta \rvert}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \geq z_{a/2}$	$Z = \dfrac{\bar{X} - \bar{Y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \geq z_a$	$Z = \dfrac{\bar{X} - \bar{Y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \leq -z_a$

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη διαφορά **$\mu_1 - \mu_2$** των μέσων τιμών δύο πληθυσμών με δύο ανεξάρτητα δείγματα μεγέθους **n_1** και **n_2** αντίστοιχα

H_0: $\mu_1 - \mu_2 = \delta$

- Οι διακυμάνσεις **σ_1^2** και **σ_2^2** είναι **άγνωστες** και τα μεγέθη των δειγμάτων **n_1** και **n_2** είναι μεγάλα **(≥30)** (οτιδήποτε πληθυσμός) [*έλεγχος επιπέδου σημαντικότητας α κατά προσέγγιση*]

Περιοχή απόρριψης της H_0

H_1: $\mu_1 - \mu_2 \neq \delta$	H_1: $\mu_1 - \mu_2 > \delta$	H_1: $\mu_1 - \mu_2 < \delta$
$\lvert Z \rvert = \dfrac{\lvert \bar{X} - \bar{Y} - \delta \rvert}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \geq z_{a/2}$	$Z = \dfrac{\bar{X} - \bar{Y} - \delta}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \geq z_a$	$Z = \dfrac{\bar{X} - \bar{Y} - \delta}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \leq -z_a$

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη διαφορά **$\mu_1 - \mu_2$** των μέσων τιμών δύο πληθυσμών με δύο ανεξάρτητα δείγματα μεγέθους **n_1** και **n_2** αντίστοιχα

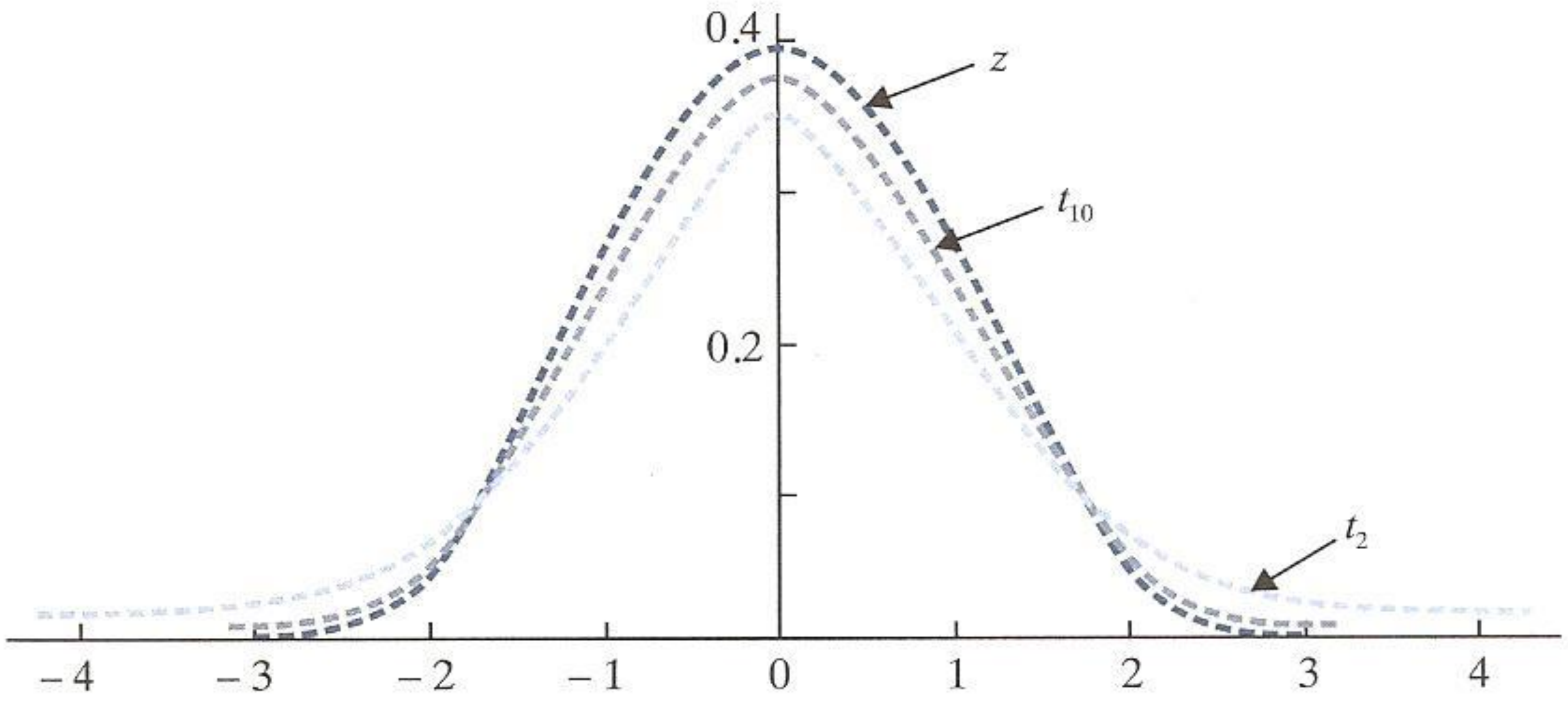
H_0: $\mu_1 - \mu_2 = \delta$

- Οι διακυμάνσεις **σ_1^2** και **σ_2^2** είναι **άγνωστες και ίσες**, οι πληθυσμοί είναι **κανονικοί** και τα μεγέθη των δειγμάτων τους **n_1** και **n_2** είναι **οτιδήποτε**
- $n = n_1 + n_2 - 2, \qquad S^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

Περιοχή απόρριψης της H_0

H_1: $\mu_1 - \mu_2 \neq \delta$	H_1: $\mu_1 - \mu_2 > \delta$	H_1: $\mu_1 - \mu_2 < \delta$
$\lvert T \rvert = \frac{\lvert \bar{X} - \bar{Y} - \delta \rvert}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \geq t_{n;a/2}$	$T = \frac{\bar{X} - \bar{Y} - \delta}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \geq t_{n;a}$	$T = \frac{\bar{X} - \bar{Y} - \delta}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \leq -t_{n;a}$

Η κατανομή *t* (ή Student) με *n* βαθμούς ελευθερίας – καμπύλη



Η συνάρτηση πυκνότητας της t_v για $v = 2$ και $v = 10$ και της $N(0, 1)$.

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη διαφορά $\mu_1 - \mu_2$ των μέσων τιμών δύο πληθυσμών με δύο ανεξάρτητα δείγματα μεγέθους n_1 και n_2 αντίστοιχα

H_0: $\mu_1 - \mu_2 = \delta$

- Όταν τα **n_1** και **n_2** είναι **μικρά**, ο πληθυσμοί **όχι κανονικοί** και οι διακυμάνσεις **σ_1^2** και **σ_2^2** είτε γνωστές είτε άγνωστες (ίσες ή όχι) δεν μπορούμε να πραγματοποιήσουμε στατιστικό έλεγχο υποθέσεων σε επίπεδο σημαντικότητας α

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη διαφορά δύο διωνυμικών ποσοστών $\mathbf{p_1}$ και $\mathbf{p_2}$ με δύο ανεξάρτητα δείγματα μεγέθους $\mathbf{n_1}$ και $\mathbf{n_2}$ αντίστοιχα

H_0: $p_1 = p_2$

- $n_i\widehat{p_i} \geq 5$ και $n_i(1-\widehat{p_i}) \geq 5$, $i = 1, 2$
- $\widehat{p_i}$ είναι το ποσοστό επιτυχιών στο δείγμα n_i, $i = 1, 2$
- $\hat{p} = \frac{n_1\widehat{p_1}+n_2\widehat{p_2}}{n_1+n_2}$

Περιοχή απόρριψης της H_0

H_1: $p_1 \neq p_2$	H_1: $p_1 > p_2$	H_1: $p_1 < p_2$
$\frac{\lvert\widehat{p_1}-\widehat{p_2}\rvert}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \geq z_{a/2}$	$\frac{(\widehat{p_1}-\widehat{p_2})}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \geq z_a$	$\frac{(\widehat{p_1}-\widehat{p_2})}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \leq -z_a$

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για την ισότητα των διακυμάνσεων $\boldsymbol{\sigma_1^2}$ και $\boldsymbol{\sigma_2^2}$ δύο **κανονικών** πληθυσμών με δύο ανεξάρτητα δείγματα μεγέθους $\mathbf{n_1}$ και $\mathbf{n_2}$ αντίστοιχα

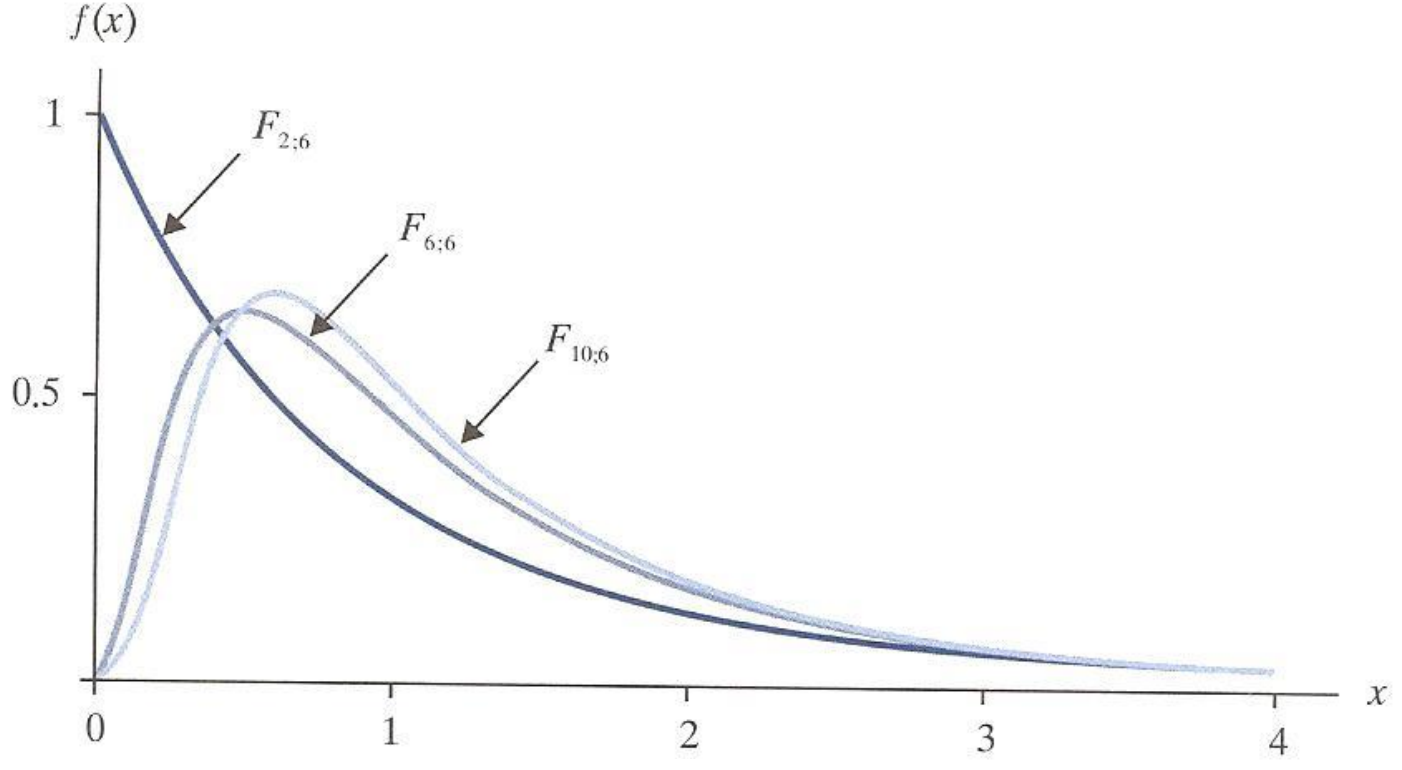
H_0: $\sigma_1^2 = \sigma_2^2$

Περιοχή απόρριψης της H_0

H_1: $\sigma_1^2 \neq \sigma_2^2$	H_1: $\sigma_1^2 > \sigma_2^2$	H_1: $\sigma_1^2 < \sigma_2^2$
$\frac{s_1^2}{s_2^2} \geq F_{n_1-1;n_2-1;a/2}$ ή $\frac{s_1^2}{s_2^2} \leq F_{n_1-1;n_2-1;1-a/2}$	$\frac{s_1^2}{s_2^2} \geq F_{n_1-1;n_2-1;\alpha}$	$\frac{s_1^2}{s_2^2} \leq F_{n_1-1;n_2-1;\alpha}$

Η τ.μ. $\frac{s_1^2}{s_2^2}$, για κανονικούς πληθυσμούς, προσεγγίζεται ικανοποιητικά από την F με n_1-1 και n_2-1 βαθμούς ελευθερίας

Η κατανομή *F* με *n* και *m* βαθμούς ελευθερίας – καμπύλη



Η συνάρτηση πυκνότητας της $F_{v;m}$ για $v = 2,6,10$ και $m = 6$.

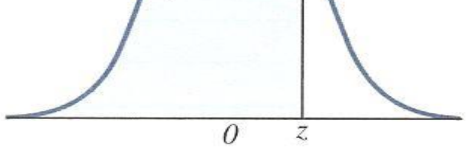
Παράδειγμα 3

- Μια αυτόματη μηχανή συσκευάζει καλαμπόκι σε τσουβάλια των 25 Kg
- Η μηχανή έχει ρυθμιστεί έτσι ώστε οι ποσότητες καλαμποκιού που συσκευάζονται ανά τσουβάλι να έχουν τυπική απόκλιση 1.5 Kg
- Επίσης, έχει παρατηρηθεί ότι οι ποσότητες αυτές ακολουθούν κανονική κατανομή
- Ο υπεύθυνος παραγωγής υποψιάζεται ότι η μηχανή έχει απορυθμιστεί και θέλει να ελέγξει αν η τυπική απόκλιση των ποσοτήτων καλαμποκιού που συσκευάζονται ανά τσουβάλι είναι πράγματι 1.5 Kg
- Για το σκοπό αυτό, από την παραγωγή μιας ημέρας, επέλεξε τυχαία 30 τσουβάλια, κατέγραψε τα βάρη τους και υπολόγισε τον μέσο τους και την τυπική τους απόκλιση και βρήκε $\bar{x} = 24.8$ Kg και $s = 1.6$ Kg
- Σε επίπεδο σημαντικότητας 5%, υποστηρίζουν τα δεδομένα αυτά τις υποψίες του υπεύθυνου παραγωγής;

ν	$\alpha = 0.995$	$\alpha = 0.99$	$\alpha = 0.975$	$\alpha = 0.95$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
1	0.000	0.000	0.001	0.004	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.647	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.414	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.878	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.994
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.335
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672
40	20.706	22.164	24.433	26.509	55.756	59.342	63.691	66.766
50	27.991	29.708	32.357	34.764	67.505	71.420	76.154	79.490
60	35.535	37.485	40.482	43.188	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.392	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.930	124.342	129.561	135.807	140.169

Παράδειγμα 4

- Η εταιρία ύδρευσης μιας πόλης ενδιαφέρεται να γνωρίζει τη συγκέντρωση μόλυβδου στο πόσιμο νερό που φτάνει μέσω του δικτύου της στους καταναλωτές και για το σκοπό αυτό κάνει τακτικούς δειγματοληπτικούς ελέγχους
- Επειδή το νερό που φτάνει στις διάφορες περιοχές της πόλης δεν προέρχεται από τους ίδιους ταμιευτήρες, όπως επίσης και το δίκτυο ύδρευσης δε βρίσκεται στην ίδια κατάσταση σε όλους τους τομείς της πόλης, οι δειγματοληπτικοί έλεγχοι γίνονται ανά πολεοδομικό τομέα
- Σε πρόσφατο δειγματοληπτικό έλεγχο, ένα δείγμα 100 παροχών από το κέντρο της πόλης έδωσε μέση συγκέντρωση μόλυβδου 36 ppm με τυπική απόκλιση 6 ppm, ενώ ένα δείγμα 90 παροχών από τα ανατολικά προάστια έδωσε μέση συγκέντρωση 34.1 ppm με τυπική απόκλιση 5.9 ppm
- Άραγε αυτή η διαφορά των 1.9 ppm, μεταξύ των δύο δειγμάτων, είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 5%;
- Δηλαδή, προέκυψε επειδή η μέση συγκέντρωση μόλυβδου στο κέντρο της πόλης πράγματι διαφέρει από τη μέση συγκέντρωση στα ανατολικά προάστια ή προέκυψε τυχαία (οφείλεται στην τύχη);



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998