Κατανομές δειγματοληψίας (Sampling distributions)

Κατανομή δειγματοληψίας (sampling distribution)

- □ Το στατιστικό στοιχείο είναι μια τυχαία μεταβλητή που περιγράφει συνολικά το δείγμα
- Η κατανομή που ακολουθεί το στατιστικό στοιχείο ονομάζεται κατανομή δειγματοληψίας και αποτελεί μία γνωστή κατανομή
- □ Μελετώντας το στατιστικό στοιχείο, μέσω της κατανομής δειγματοληψίας, μπορούμε να κάνουμε:
 - ✓ εκτίμηση της τιμής της παραμέτρου
 - ✓ έλεγχο κάποιας υπόθεσης πάνω στην παράμετρο

Σημαντικά στατιστικά στοιχεία

- \Box Ο δειγματικός μέσος (sample mean) $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
- □ Η δειγματική διακύμανση (sample variance)

$$S^{2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i} - \overline{X})^{2}$$

Ισχύει η σχέση:
$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} X_i^2 - \frac{n}{n-1} \overline{X}^2$$

S (>0) είναι η δειγματική τυπική απόκλιση

- □ Άλλα στατιστικά στοιχεία είναι:
 - το μέγιστο (max), το ελάχιστο(min),
 - το εύρος τιμών (/max-min/),
 - кλπ.

Σπουδαιότερες κατανομές δειγματοληψίας

- Η Κανονική κατανομή Ν(μ, σ²)
- Η κατανομή *Χι-τετράγωνο* X_n^2
- Η κατανομή t-Student t_n
- Η F-κατανομή [F_{n,m}]

(A). Η Κανονική κατανομή *Ν(μ, σ*²)

 $x \sim N(\mu, \sigma^2)$ πυποποίησης $z = \frac{x - \mu}{\sigma}$ $z = \frac{x - \mu}{\sigma}$

$$X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

(τυπική κανονική)

Πρόταση 1α

• Αν X_1 , X_2 , ..., X_n \mathbf{n} ανεξάρτητες και ισόνομες κανονικές μεταβλητές, δηλ. $X_i \sim N(\mu, \sigma^2)$, τότε για το άθροισμα ισχύει ότι:

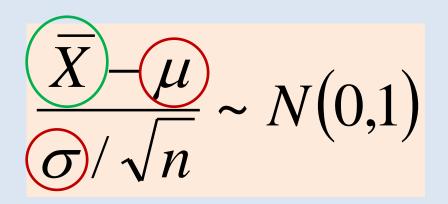
$$S_n = \sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2) \quad Z = \frac{S_n - n\mu}{\sqrt{n}\sigma} \sim N(0,1)$$

 Παρατήρηση: Από το Κ.Ο.Θ. το παραπάνω ισχύει για κάθε κατανομή όταν το n είναι μεγάλο.

Πρόταση 1β

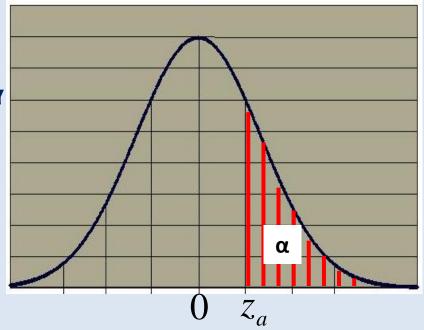
• Αν X_1 , X_2 , ..., X_n n ανεξάρτητες και ισόνομες κανονικές μεταβλητές, δηλ. $X_i \sim N(\mu, \sigma^2)$, τότε για τον δειγματικό μέσο ισχύει ότι:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_{i}}{n} \sim N\left(\mu, \frac{\sigma^{2}}{n}\right) \Rightarrow \overline{Z} = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$



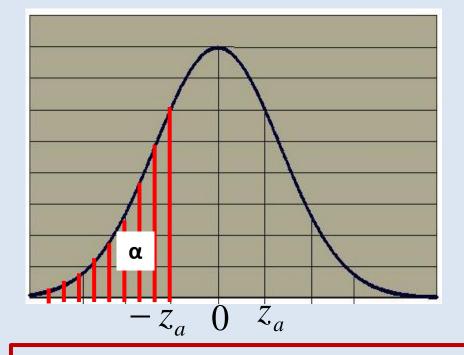
- Το στατιστικό στοιχείο του δειγματικού μέσου (ή μέσου όρου) σχετίζεται με τις παραμέτρους του μέσου μ και της διασποράς σ²
- Η κατανομή αυτού του στοιχείου είναι η τυπική κανονική (με μέσο 0 και διακύμανση 1)

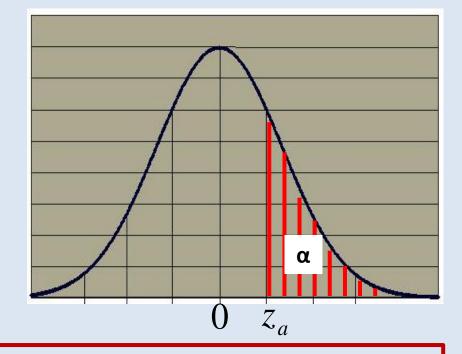
Πρόταση 2: Το σημείο Ζα



 Ορίζουμε ως z_a εκείνο το σημείο εκείνο για το οποίο ισχύει :

$$z_a: P(X > z_a) = a$$



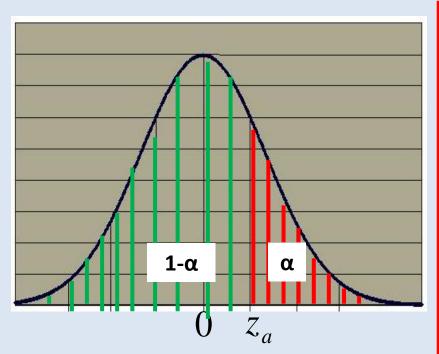


 Ορίζουμε ως z_a εκείνο το σημείο εκείνο για το οποίο ισχύει (λόγω συμμετρίας):

$$z_a: P(X > z_a) = P(X \le -z_a) = a$$

Διαδικασία εύρεσης του σημείου Ζα

 Εύρεση από τον πίνακα του σημείου που αντιστοιχεί στην πιθανότητα 1-α (στοιχείο πίνακα)

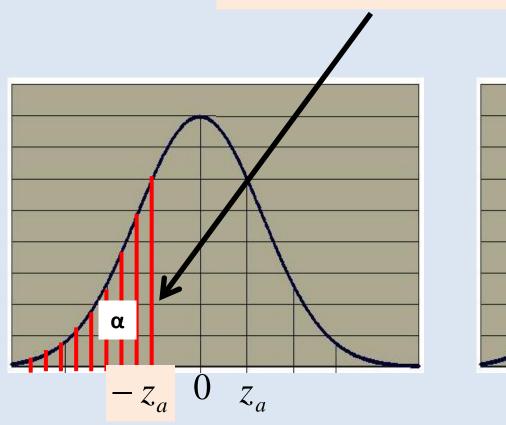


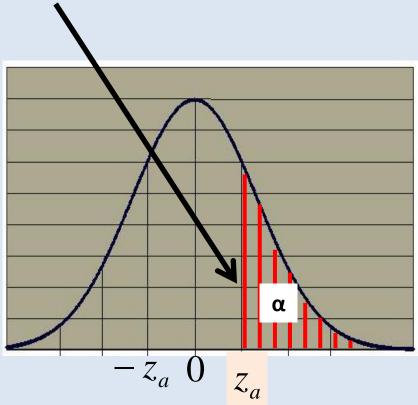
							/::	\mathbb{R}^{n}		
							/::::	::: N		
							Carania de		/	
Æ()	$=\int_{-1}^{1}$	^r 1	t2	12			-:-:-:	-:-:-	-	_
$\Psi(x)$	- /	/0	_ e ·	$'$ - $a\iota$:	mmmi	minimi	ininini n	inini mu		minimum.
	J –	∞ V 2	и		-3	-2	-1	0 1		3
m-1-	.1 1				C . 1	1.707	1.0	x urve to		e. e
Tac										
x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.0	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9982	.9982	.9988	.9988	.9989	.9989	.9989	.9990	.9990
	.9987	.9987	.9987	.9998		.9989	.9989		.9993	.9993
3.1					.9992			.9992		
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Έτσι

$$z_a$$
: $\Phi(z_a) = 1 - a$

$$a = \Phi(-z_a) = 1 - \Phi(z_a)$$

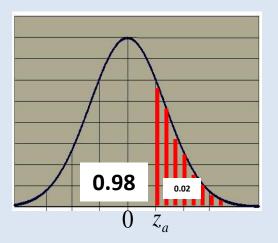




Παράδειγμα: Να βρεθεί το σημείο **Ζ**_{0.02}

$$z_a: \Phi(z_a) = 1 - a = 0.98$$

$$z_a = 2.05$$



Tol	do E 1	A ==== #	(m)	dan tha	Standa	and No.		x	the le	ft of x.
x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
$\frac{3.1}{3.2}$.9990	.9991	.9991	.9991	.9992 .9994	.9992	.9992	.9992	.9993	.9993
3.3	.9995	.9995	.9994	.9994	.9994	.9994	.9994	.9996	.9995	.9995
	.9995	.9995		.9996		.9997		.9996	.9996	
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Η κατανομή **Χι-τετράγωνο Χ²**_n με **n** βαθμούς ελευθερίας

• Η κατανομή **Γάμμα** *G*(*a*, *λ*) έχει σ.π.π.

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \qquad x \ge 0$$

Αν α=n/2 και λ=1/2 τότε:

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

προκύπτει η συνάρτηση της κατανομής **Χι-τετράγωνο** με **n** βαθμούς ελευθερίας, **Χ**²_n

Η μορφή της συνάρτησης πυκνότητας

:



Πρόταση f(x) $v^{2}(a)$ $\chi^{2}(a)$ $\chi^{2}(a)$

• Ορίζουμε ως χ^2_n (a) εκείνο το σημείο εκείνο για το οποίο ισχύει:

$$P(X > \chi_n^2(a)) = a$$

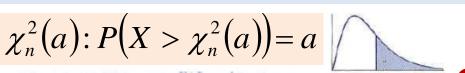
$$P(X > \chi) = \int_{\chi}^{\infty} f(x) dx = \int_{\chi}^{\infty} \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2} dx$$

όχι αναλυτικός τύπος (αριθμητικές μέθοδοι υπολογισμού του ολοκληρώματος)

Πίνακας τιμών της χι-τετράγωνο κατανομής:

για διάφορες τιμές της πιθανότητας α και του βαθμού ελευθερίας η





 $P(X > \chi) = \int_{\gamma}^{\infty} \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2 - 1} e^{-x/2} dx$ f(x)

	<i>i</i> ()			2011	•			7	_		,				χ_n^2	(a)				\mathcal{X}
T	able A.5 (Critical Va	lues of the	Chi-Squar	ed Distribu	ition	0	10 No.							70 <u>n</u>	.\/_				_
					α						_					α				
v	0.995	0.99	0.98	0.975	0.95	0.90	0.80	0.75	0.70	0.50	0.3	0.25	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.001
1 2	0.0^4393 0.0100	0.0^3157 0.0201	0.0^3628 0.0404	0.03982	0.00393	0.0158	0.0642	0.102	0.148	0.455 1.386	1.074 2.408	2,773	1.642 3.219	2.706 4.605	3.841 5.991	5.024 7.378	5.412 7.824	6.635 9.210	7.879	10.827 13.815
3	0.0717	0.115	0.185	0.216	0.103	0.584	1.005	1.213	1.424	2.366	3.665	4.108	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.266
4	0.207	0.297	0.429	0.484	0.711	1.064	1.649	1.923	2.195	3.357	4.878	5.385	5.989	7.779	9.488	11.143	11.668	13.277	14.860	18.466
5	0.412	0.554	0.752	0.831	1.145	1.610	2.343	2.675	3.000	4.351	6.064	6.626	7.289	9.236	11.070	12.832	13.388	15.086	16.750	20.515
6	0.676	0.872	1.134	1.237	1.635	2.204	3.070	3.455	3.828	5.348	7.231	7.841	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457
7	0.989	1.239	1.561	1.690	2.167	2.833	3.822	4.255	4.671	6.346	8.383	9.037	9.803	12.017	14.067	16.013	16.622	18.475	20.278	24.321
8	1.344	1.647	2.032	2.1.80	2.733	3.490	4.594	5.071	5.527	7.344	9.524	10.219	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.124
9	1.735	2.088	2.532	2.700	3.325	4.168	5.380	5.899	6.393		10.656	11.389	12.242	14.684	16.919	19.023	19.679	21.666	23.589	27.877
10	2.1.56	2.558	3.059	3.247	3.940	4.865	6.179	6.737	7.267	9.342	11.781	12.549	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588
11	2.603	3.053	3.609	3.816	4.575	5.578	6.989	7.584	8.148	10.341		13.701	14.631	17.275	19.675	21.920	22.618	24.725	26.757	31.264
12	3.074	3.571	4.178	4.404	5.226	6.304	7.807	8.438	9.031	11.340		14.845	15.812	18.549	21.026	23.337	24.054	26.217	28.300	32.909
13	3.565	4.107	4.765	5.009	5.892	7.041	8.634	9.299	9.926	12.340		15.984	16.985	19.812	22.362	24.736	25.471	27.688	29.819	34.527
14	4.075	4.660	5.368	5.629	6.571	7.790	9.467	10.165	10.821			17.117 18.245	18.151 19.311	21.064	23.685	26.119	26.873 28.259	29.141 30.578	31.319	36.124 37.698
15	4.601	5.229	5.985	6.262	7.261	8.547	10.307	11.037	11.721	14.339				2	10	\				
16	5.142	5.812	6.614	6.908	7.962	9.312	11.152	11.912		15.338		19.369 20.489	20.465 21.615	V	(0.2)	25 I	29.633 30.995	32.000	34.267 35.718	39.252 40.791
17 18	5.697 6.265	6.408 7.015	7.255 7.906	7.564 8.231	8.672 9.390	10.085	12.002 12.857	12.792	13.531			21.605	22.760	$\lambda 20$	(0.4	<u> </u>	32.346	34.805	37.156	42.312
10	6.844	7.633	8.567	8.907	10.117	11,651	13.716		15.352			22.713	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.819
20	.434	8.260	9.237	9.591	10.851	12.443	14.578		16.266			23.828		28.412	31.410	34.170	35.020	37.566	39.997	45.314
21	8.034	8.897	9.915	10.283	11.591	13.240	15.445	16.344	17.182	20.337	13.858	24.955	26.171	29.615	32.671	35.479	36.343	38.932	41.401	46.796
22	8.643	9.542	10.600	10.982	12.338	14.041	16.314	17.240		21.337		26.039	27.301	30.813	33.924	36.781	37.659	40.289	42.796	48.268
23	9.260	10.196	11.293	11.689	13.091	14.848	17.187	18.137	19.021	22.337	16.018	27.141	28.429	32.007	35.172	38.076	38.968	41.638	44.181	49.728
24	9.886	10.856	11.992	12.401	13.848	15.659	18.062	19.037		23.337		28.241	29.553	33.196	36.415	39.364	40.270	42.980	45.558	51.179
25	10.520	11.524	12.697	13.120	14.611	16.473	18.940	19.939	20.867	24.337	8.172	29.339	30.675	34.382	37.652	40.646	41.566	44.314	46.928	52.619
20	11.160	12.198	13.409	13.844	15.379	17.292	19.820	20.843	21.792	23.330	19.246	30.435	31.795	35.563	38.885	41.923	42.856	45.642	48.290	54.051
27	11.808	12.878	14.125	14.573	16.151	"1.8.114	20.703	21.749	22.719			31.528	32.912	36.741	40.113	43.195	44.140	46.963	49.645	55.475
28	12.461	13.565	14.847	15.308	16.928	18.939	21.588	22.657	23.647			32.620	34.027	37.916	41.337	44.461	45.419	48.278	50.994	56.892 58.301
30	13.121	11.256	15.574	16.047	17.708	19.768	22.475	23.567	24.577	28.336		33.711	35.139 36.250	39.087 40.256	42.557	45.722 46.979	46.693 47.962	49.588 50.892	52.335 53.672	59.702
	13.787		16.306	16.791	18.493	20.599	23.364	24.478										63.691	66.766	73.403
40 50	20.707 27.991	22.164	23.838	24.433	26.509 34.764	29.051	32.345	33.66	34.872	39.335		45.616 56.334	47.269 58.164	51.805 63.167	55.758 67.505	59.342 71.420	60.436 72.613	76.154	79.490	73.403 86.660
	35.534	37.485	31.664 39.699	32.357 40.482	43.188	37.689 46.459	41.449 50.641		44.313 53.809			66.981	68.972	74.397	79.082	83.298	84.58	88.379	91.952	99.608
		27.103	22.000	311/11/16	.5.100	40.435	50.041	22.224	23.009	39.333			-			20.000				

Πρόταση

- Έστω *n* τ.μ. {X₁, X₂, ..., X_n} ανεξάρτητες και ισόνομες που ακολουθούν την τυπική κανονική, δηλ. X_i ~ N(0, 1).
- Τότε η $Y = \sum_{i=1}^n X_i^2$ ακολουθεί την κατανομή **Χι-τετράγωνο**

με n βαθμούς ελευθερίας X_{n}^{2} .

$${X_i \sim N(0,1)}_{i=1}^n \Longrightarrow Y = \sum_{i=1}^n X_i^2 \sim X_n^2$$

με συνάρτηση πυκνότητας πιθανότητας:

$$f(y) = \frac{1}{2^{n/2} \Gamma(n/2)} y^{n/2-1} e^{-y/2}$$

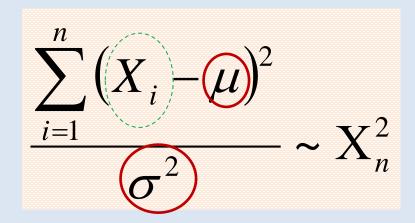
Εφαρμογές της κατανομής Χι-τετράγωνο στη Στατιστική

Πρόταση 1

• Το στατιστικό στοιχείο
$$\frac{\displaystyle\sum_{i=1}^{n}(X_{i}-\mu)^{2}}{\sigma^{2}}$$
 ακολουθεί την

κατανομή Χι-τετράγωνο με *n* βαθμούς ελευθερίας

$$\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim X_n^2$$



- Το στατιστικό στοιχείο συνδέει το δείγμα με τις παραμέτρους του μέσου μ και της διασποράς σ²
- Η κατανομή που ακολουθεί είναι η Χι-τετράγωνο με *n* βαθμούς ελευθερίας.

Πρόταση 2

• Το στατιστικό στοιχείο
$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2}$$

ακολουθεί την κατανομή Χι-τετράγωνο με η-1 βαθμούς ελευθερίας

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2} \sim X_{n-1}^2$$

• όπου $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$ η δειγματική διασπορά

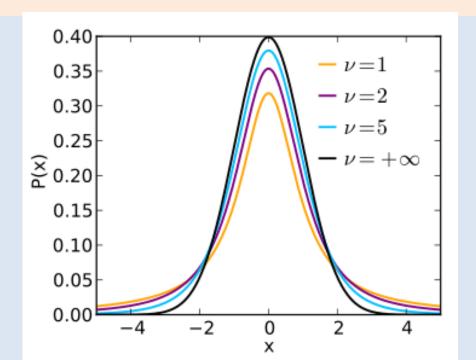
$$\frac{\sum_{i=1}^{n} (X_{i} - \overline{X})^{2}}{(\sigma^{2})} \sim X_{n-1}^{2} \Rightarrow \frac{(n-1)S^{2}}{(\sigma^{2})} \sim X_{n-1}^{2}$$

- Αυτό το στατιστικό στοιχείο συνδέει το δείγμα με την παράμετρο της διασποράς σ²
- Η κατανομή που ακολουθεί είναι η Χι-τετράγωνο με *n-1* βαθμούς ελευθερίας

(Γ). Η κατανομή t-Student (Gosset 1908)

Συνάρτηση πυκνότητας πιθανότητας (σ.π.π.)

$$f_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\sqrt{n}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \frac{1}{\sqrt{\pi}B\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

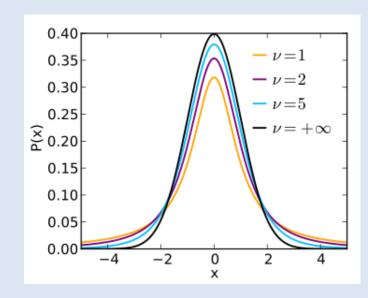


$$B(a+b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Ιδιότητες της t-Student κατανομής

• Μέση τιμή και διακύμανση:

$$E(X)=0$$
 KQI $V(X)=n/(n-2)$



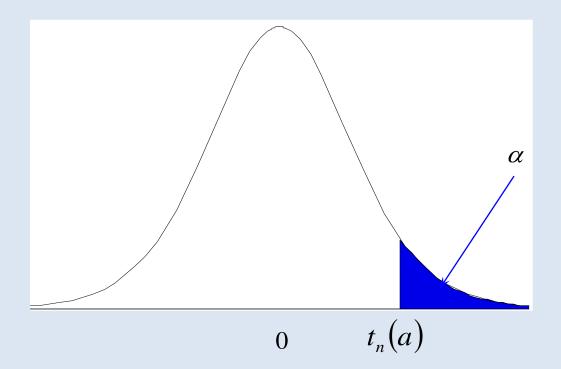
• Είναι συμμετρική

μοιάζει με τη Κανονική κατανομή, αλλά έχει πιο πολύ πιθανότητα στις άκρες.

• Για μεγάλη τιμή του 'n' συγκλίνει στην τυπική Κανονική

$$\lim_{n\to\infty} f_n(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

Πρόταση



Ορίζουμε ως t_n(\alpha) το σημείο για το οποίο ισχύει:

$$P(X > t_n(a)) = a$$

π.χ. εύρεση του t_{12} (0.005)

			Total A Dis		_				t_n	(a):	P(x)	X >	$rac{t_n}{(}$	(a)	= a	\mathcal{I}
Table	A.4 Criti	cal Values	of the 1-1719				-	a					α			
ν	0.40	0.30	0.20	α 0.15	0.10	0.05	0.025		v	0.02	0.015	0.01	0.0075	0.005	0.0025	0.
1				1.963					1	15.894	21.205	31.821	42.433	53.656	127.321	63
2	0.325	0.727	1.376 1.061	1.386	3.078 1.886	6.314 2.920	12.706 4.303		2	4.849	5.643	6.965	8.073	9.925	14.089	3
3	0.239	0.584	0.978	1.250	1.638	2.353	3.182		3	3.482	3.896	4.541	5.047	5.841	7.453	
a	0.277	0.569	0.941	1.190	1.533	2.333	2.776		4	2.999	3.298	3.747	4.088	4.604	5.598	
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571		5	2.757	3.003	3.365	3.634	4.032	4.773	
									6	2.612	2.829	3.143	3.372	3.707	4.317	
6 7	0.265	0.553	0.906	1.134	1.440 1.415	1.943	2.447		7	2.517	2.715	2.998	3.203	3.499	4.029	
8	0.263	0.549	0.896	1.119		1.895			8	2.449	2.634	2.896	3.085	3.355	3.833	
i)	0.262	0.546	0.889	1.108	1.397	1.860 1.833	2.306		9	2.398	2.574	2.821	2.998	3.250	3.690	
:10	0.260	0.543	0.883	1.003	1.372	1.833	2.202		10	2.359	2.527	2.764	2.932	3.169	3.581	
110	0.260	0.540	0.879	1.093	1.372	1.706	2.228		11	2.328	2.491	2.718	2.879	1.106	3.497	
2000									12	2.303	2.461	2.681	2.836	3.055	3.428	
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179		13	2.303	2.436	2.650	2.801	3.012	3 372	
111	0.258	0.537	0.868	1.076	1.345	1.761	2.145		14	2.264	2.415	2.624	2.771	2.211	3.326	
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131		15	2.249	2.397	2.602	2.746	2.947	3.286	
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120		16	2.235	2.382	2.583	2.724	2.921	3.252	
:17	0.257	0.534	0.863	1.069	1.333	1.740	2.110		17	2.224	2.368	2.567	2:706	2.898	3.222	
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101		18	2.214	2.356	2.552	2.689	2.878	3.197	
15	0.257	0.533	0.861	1.066	1.328	1.729	2.093		19	2.205	2.346	2.539	2.674	2.861	3.174	
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086		20	2.197	2.336	2.528	2.661	2.845	3.153	
21	0.257	0.532	0.859	1.063	1.323	1.721			21	2.189	2.328	2.518	2.649	2.831	3.135	
22	0.256	0.532	0.858	1.061	1.323	1.721	2.080		22	2.183	2.320	2.508	2.639	2.831	3.119	
23	0.256	0.532	0.858	1.060	1.321	1.717	2.069		23	2.177	2.320	2.500	2.629	2.819	3.119	
24	0.256	0.531	0.857	1.059	1.319	1.714	2.064		24	2.177	2.307	2.492	2.629	2.797	3.091	
25	0.256	0.531	0.856	1.058	1.316	1.705	2.060		25	2.172	2.301	2.485	2.612	2.787	3.078	
		0.531														
26 27	0.256	0.531	0.856	1.058	1.315	1.706	2.056		26	2.162	2.296	2.479	2.605	2.779	3.067	
28	0.256	0.531	0.855	1.056		1.703	2.052		27	2.158	2.291	2.473	2.598	2.771	3.057	
29	0.256	0.530	0.854	1.055	1.313	1.701	2.048		28	2.154	2.286	2.467	2.592	2.763	3.047	
30	0.256	0.530	0.854	1.055	1.311	1.697	2.043		29	2.150	2.282	2.462	2.586	2.756	3.038	
									30	2.147	2.278	2.457	2.581	2.750	3.030	
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021		40	2.123	2.250	2.423	2.542	2.704	2.971	
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000		60	2.099	2.223	2.390	2.504	2.660	2.915	
120	0.254	0.526 0.524	0.845 0.842	1.041	1.289	1.658	1.980		120	2.076	2.196	2.358	2.468	2.617	2.860	
00	0.233	0.324	0.042	1.030	.707	1.645	1.960	.*	- ∞	2.054	2.170	2.326	2.432	2.576	2.807	

 $t_{12}(0.005) = 3.055$

π.χ. εύρεση του t_{11} (0.05)

Tabl	le A.4 Criti	cal Values	of the I-Di	erdilution	_		-		t_1	n(a)): <i>P</i>	(X	$> t_{\eta}$	n(a)))=	a
-			74 1	α									α			
ν	0.40	0.30	0.20	0.15	0.10	0.05	0.025		V	0.02	0.015	0.01	0.0075	0.005	0.0025	0.0005
1	0.325	0.727	1.376	1.963	3.078	-37	12.706		1	15.894	21.205	31.821	42.433	63.656	127.321	636.578
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303		2	4.849	5.643	6.965	8.073	9.925	14.089	31.600
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182		3	3.482	3.896	4.541	5.047	5.841	7.453	12.924
(1	0.271	0.569	0.941	1.190	1.533	2.132	2.776		4	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571		5	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447		6	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365		7	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306		8	2.449	2.634	2.896	3.085	3.355	3.833	5.041
i)	0.261	0.543	0.883	1.100	1.383		2.262		9	2.398	2.574	2.821	2.998	3.250	3.690	4.781
	0.260	0.542	0.870	1.003	1.372	1.812	2 220		10	2.359	2.527	2.764	2.932	3.169	3.581	4.587
3.1	0.260	0.540	0.876	1.088	1.363	1.796	2.201		11	2.328	2.491	2.718 2.681	2.879	3.106	3.497 3.428	4.437
10	0.259	0.539	0.873	1.083	1.350	1.782	2.179		12	2.303	2.436	2.650	2.898	3.012	3.372	4.221
13	0.259	0.538	0.870	1.079	1.350		2.1 👀		14	2.262	2.436	2.624	2.771	2.977	3.326	4.221
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145		15	2.249	2.415	2.602	2.771	2.947	3.286	4.073
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131									
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120		16	2.235	2.382	2.583	2.724	2.921	3.252 3.222	4.015
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110		17	2.224	2.368	2.567	2.706	2.898		3.965 3.922
1.8	0.257	0.534	0.862	1.067	1.330	1.734	2:101 2:093		18 19	2.214	2.356	2.552 2.539	2.689 2.674	2.878	3.197 3.174	3.883
15) 20	0.257	0.533	0.861	1.066	1.328	1.729				2.205	2.346			2.861	3.174	3.883
	0.257	0.533	0.860	1.064	1.325	1.725	2.086		20	2.197	2.336	2.528	2.661	2.845		
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080		21	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074		22	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069		23	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24 25	0.256	0.531	0.857	1.059	1.318	1.711	2.064		24	2.172	2.307	2.492	2.620	2.797	3.091	3.745
	0.256	0.531	0.856	1.058	1.316	1.70S	2.060	<u>.</u>	25	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056	s s	26	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052	8	27	2.158	2.291	2.473	2.598	2.771	3.057	3.689
28 29	0.256	0.530	0.855	1.056	1.313	1.701	2.048	# #	28	2.154	2.286	2.467	2.592	2.763	3.047	3.674
30	0.256 0.256	0.530	0.854 0.854	1.055	1.311	1.699	2.045 2.042	18 18	29	2.150	2.282	2.462	2.586	2.756	3.038	3.660
						1.697		*	30	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021	18 18	40	2.123	2.250	2.423	2.542	2.704	2.971	3.551
60 120	0.254 0.254	0.527	0.848	1.045	1.296	1.671	2.000	**	60	2.099	2.223	2.390	2.504	2.660	2.915	3.460
∞	0.254	0.526 0.524	0.845 0.842	1.041	1.289	1.658 1.645	1.980 1.960	**	120	2.076	2.196	2.358	2.468	2.617	2.860	3.373
	0.233	0.324	0.072	1.030	.200	1.043	1.900	*	- 00	2.054	2.170	2.326	2.432	2.576	2.807	3.290

 $t_{11}(0.05) = 1.796$

$$\frac{a}{2} = P\left(T < -t_n\left(\frac{a}{2}\right)\right)$$

$$1 - a$$

$$-t_n\left(\frac{a}{2}\right)$$

$$0$$

$$t_n\left(\frac{a}{2}\right)$$

$$P\left(-t_{n}\left(\frac{a}{2}\right) < T < t_{n}\left(\frac{a}{2}\right)\right) = P\left(T < t_{n}\left(\frac{a}{2}\right)\right) - P\left(T < -t_{n}\left(\frac{a}{2}\right)\right) = P\left(T > -t_{n}\left(\frac{a}{2}\right)\right) - P\left(T > t_{n}\left(\frac{a}{2}\right)\right) = 1 - 2P\left(T > t_{n}\left(\frac{a}{2}\right)\right)$$

Πρόταση 3

• Έστω 2 ανεξάρτητες τυχαίες **Ζ, Υ** μεταβλητές με κατανομές:

Z~N(0,1) και **Y~ X_{n}^{2}**, τότε η τυχαία μεταβλητή $T = \frac{Z}{\sqrt{Y/n}}$

$$T = \frac{Z}{\sqrt{Y/n}}$$

ακολουθεί την κατανομή *t-Student* με *n* βαθμούς ελευθερίας.

$$Z \sim N(0,1) \Longrightarrow T = \frac{Z}{\sqrt{Y/n}} \sim t_n$$

Εφαρμογή της Πρότασης 3

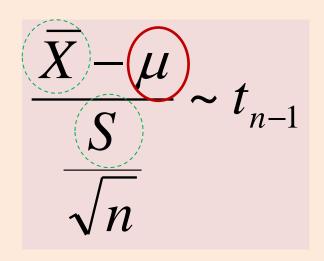
Για
$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$
 και $Y = \frac{(n-1)S^2}{\sigma^2} \sim X_{n-1}^2$

έχουμε:

$$T = \frac{Z}{\sqrt{\frac{Y}{n-1}}} = \frac{\frac{\overline{X} - \mu}{\sigma / \sqrt{n}}}{\sqrt{\frac{(n-1)S^{2}}{\sigma^{2}(n-1)}}} = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

Κατανομή t-Student με n-1 βαθμούς ελευθερίας

• Αυτό το **στατιστικό στοιχείο** συνδέει το δείγμα με τη **παράμετρο** του μέσου *μ* του πληθυσμού



 Η κατανομή που ακολουθεί είναι η t-Student με n-1 βαθμούς ελευθερίας.

(Δ). H K α T α VO μ $\hat{\eta}$ F (Fisher – Snedecor 1934)



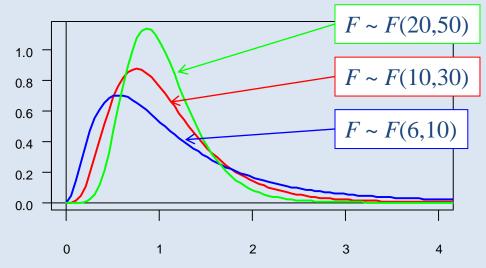
Ronald Fisher (1880– 1962), a pioneer in statistics George Snedecor (1881–1974)

Αν $Y_1 \sim X_n^2$ και $Y_2 \sim X_m^2$ τότε το πηλίκο

$$F = \frac{Y_1/n}{Y_2/m} \sim F_{n,m}$$

ακολουθεί την κατανομή F με συνάρτηση πυκνότητας:

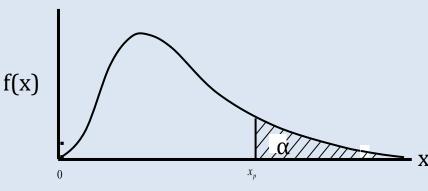
$$f_{n,m}(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \frac{x^{n/2-1}}{(1+x)^{(n+m)/2}}$$



Ιδιότητες της κατανομής F

- Η κατανομή εξαρτάται από τις τιμές των n, m αλλά και από την σειρά με την οποία εμφανίζονται, δηλ. F(n,m) ≠ F(m,n)
- Παρουσιάζει **θετική ασυμμετρία** η οποία μειώνεται όσο αυξάνονται οι βαθμοί ελευθερίας *n*, *m*.
- Μέση τιμή & διακύμανση:

$$E(X) = \frac{m}{m-2} \qquad V(X) = \frac{2m(n+m-2)}{n(m-2)^2(m-4)}$$



Ορίζουμε ως $f_a(v_1, v_2)$ το σημείο για το οποίο ισχύει:

$$P(F > x_a(v_1, v_2)) = a$$

Ισχύει ότι:

$$f_{1-\alpha}(v_1, v_2) = \frac{1}{f_{\alpha}(v_2, v_1)}$$

για συγκεκριμένο α=0.05

Table A.6* Critical Values of the F-Distribution

	$f_{0.05}(v_1,v_2)$ V ₂													
v_2	1	2	3	4	5	6	7	8	9					
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54					
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38					
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81					
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00					
5	6.61	5.79	5,11	5.19	5.05	4.95	4.88	4.82	4.77					
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10					
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68					
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39					
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18					
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02					
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90					
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80					
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71					
. 14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65					
1 15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2,59					
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2,59	2,54					
17	4.45	3.59	3.20	2.96	2.81	2.70	2,61	2,55	2.49					
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2,51	2.46					
19	4.38	3.52	3.13	2.90	2.74	2.63	2.5-1	2.48	2,12					
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39					
21	4.32	3.47	3.07	2.84	2.68	2,57	2.49	2.42	2.37					
22	4.30	3.44	3.05	2.82	2.66	2,55	2.46	2.40	2.34					
23	4.28	3.42	3.03	2.80	2.64	2,53	2.44	2.37	2.32					
24	4.26	3.40	3.01	2.78	2.62	2,51	2.42	2.36	2.30					
25	4.21	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28					
26	4.23	3.37	2.98	2.74	2,59	2.47	2.39	2.32	2.27					
27	4.21	3.35	2.96	2.73	2,57	2.46	2.37	2.31	2.25					
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24					
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22					
30	4.17	3.32	2.92	2.69	2,53	2.42	2.33	2.27	2.21					
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12					
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04					
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96					
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88					

Πρόταση 5

• Έστω 2 ανεξάρτητα δείγματα μεγέθους n_1 , n_2 με δειγματικές διασπορές S^2 και S^2 αντίστοιχα Τότε:

$$A_{1,2} = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2} \sim F_{n_1 - 1, n_2 - 1}$$

όπου σ_1^2 , σ_2^2 οι πραγματικές διασπορές των δύο πληθυσμών.

Απόδειξη

Γνωρίζουμε ότι η μεταβλητή $Y = \frac{(n-1)S^2}{\sigma^2} \sim X_{n-1}^2$

Έτσι έχουμε ότι
$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim X_{n_1-1}^2$$
 και $\frac{(n_2-1)S_2^2}{\sigma_2^2} \sim X_{n_2-1}^2$

Επομένως σύμφωνα με τον ορισμό της κατανομής F έχουμε:

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}/n_1-1}{\frac{(n_2-1)S_2^2}{\sigma_2^2}/n_2-1} \sim F_{n_1-1,n_2-1} \Rightarrow \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}$$

$$\frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2} \sim F_{n_1 - 1, n_2 - 1}$$

- Αυτό το στατιστικό στοιχείο συνδέει τα δύο διαθέσιμα δείγματα με τον λόγο των παραμέτρων της διασποράς των δύο πληθυσμών σ_1^2 προς σ_2^2
- Η κατανομή που ακολουθεί είναι η *F_{n1-1, n2-1}*

Σημαντικά στατιστικά στοιχεία και κατανομές τους

$$\sqrt{\frac{\overline{X} - \mu}{\sigma / \sqrt{n}}} \sim N(0,1)$$

$$\sum_{i=1}^{\infty} \frac{\left(X_{i} - \mu\right)^{2}}{\sigma^{2}} \sim \left(X_{n}^{2}\right) \quad \frac{(n-1)S^{2}}{\sigma^{2}} \sim \left(X_{n-1}^{2}\right)$$

$$\sqrt{\frac{X}{S} - \mu} \sim t_{n-1}$$

$$\int \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1 - 1, n_2 - 1}$$

Ως προς το είδος της κατανομής δειγματοληψίας

Σημαντικά στατιστικά στοιχεία και κατανομές τους

$$\sqrt{\frac{\overline{X} - \mu}{\sigma / \sqrt{n}}} \sim N(0,1)$$

$$\frac{\overline{X} - \mu}{S / \sqrt{n}} \sim t_{n-1}$$

$$\frac{\sum_{i=1} (X_i + \mu)^2}{\sigma^2} \sim X_n^2$$

$$\frac{(n-1)S^2}{\sigma^2} \sim X_{n-1}^2$$

$$\checkmark \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}$$

Ως προς το είδος των **παραμέτρων** που ελέγχουν