ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

ΔΙΑΛΕΞΗ: ΘΕΩΡΗΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ

Διδάσκουσα: Ε. Γάκη, Επίκ. Καθηγήτρια

MBA Χίος
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
διοικώ σημαίνει αποφασίζω

Περιεχόμενα Μαθήματος

- Θεωρητικές Κατανομές
  - Διωνυμική Κατανομή
  - Poisson Κατανομή
  - Κανονική Κατανομή
  - X^2-test
  - Κατανομή Student
  - Κατανομή F

Τυχαία Μεταβλητή & Κατανομή

Τυχαία Μεταβλητή

- Η έννοια της τυχαίας μεταβλητής χρησιμοποιείται για να περιγράψει τα αποτελέσματα ενός τυχαίου πειράματος

Κατανομή

- Η έννοια της κατανομής χρησιμοποιείται για τον υπολογισμό των πιθανοτήτων των αποτελεσμάτων αυτών.

Και οι δύο έννοιες αποτελούν την βάση για την δημιουργία μοντέλων που περιγράφουν τυχαία πειράματα χωρίς να είμαστε αναγκασμένοι να ανατρέχουμε στην θεωρία πιθανοτήτων.

Θεωρητικές Κατανομές

Κάθε τυχαία μεταβλητή έχει τη δική της κατανομή πιθανότητας αλλά σε πολλές περιπτώσεις οι κατανομές αυτές έχουν μεγάλες ομοιότητες. Μπορούμε λοιπόν να δημιουργήσουμε κάποιες βασικές μορφές κατανομών τις οποίες θα χρησιμοποιούμε αντί των πραγματικών κατανομών.

Διωνυμική Κατανομή [1]

Ας θεωρήσουμε ένα πείραμα που τα αποτελέσματα του μπορεί να είναι μόνο δύο ενδεχόμενα: **επιτυχία** (E) και **αποτυχία** (A).

Το πείραμα <u>επαναλαμβάνεται πολλές φορές</u>

Σε τέτοια προβλήματα ενδιαφερόμαστε πιο πολύ για τον αριθμό επιτυχιών στο σύνολο των δοκιμών του πειράματος παρά για το ακριβές αποτέλεσμα κάθε δοκιμής

Συμβολίζουμε n τον αριθμό των δοκιμών και με X το συνολικό αριθμό των επιτυχιών.

Σκοπός να βρούμε την κατανομή (το μοντέλο) της διακριτής τ.μ. X της οποίας οι δυνατές τιμές είναι 0,1,2,…,n.

Παραδοχές:

(α) Η πιθανότητα επιτυχίας ή αποτυχίας παραμένει σταθερή σε όλες τις δοκιμές του πειράματος

(β) Οι δοκιμές του πειράματος είναι στοχαστικά ανεξάρτητες.

Η πιθανότητα επιτυχίας είναι p, η πιθανότητα αποτυχίας είναι q, $P(E)=p$, $P(A)=q$, $p+q=1$

Διωνυμική Κατανομή [2]

Ένα πείραμα το οποίο αποτελείται από δοκιμές η κάθε μια από τις οποίες έχει μόνο δύο δυνατά αποτελέσματα Ε ή Α και για το οποίο ισχύουν οι παραπάνω παραδοχές (α) και (β) λέγεται *διωνυμικό πείραμα*. (Το διωνυμικό πείραμα λέγεται επίσης και πείραμα επαναληπτικών δοκιμών)

Δοκιμή: Κάθε επιμέρους δοκιμή λέγεται και δοκιμή Bernoulli.

Δειγματικός Χώρος: Ο δ.χ. με n δοκιμές περιλαμβάνει 2^n σημεία. Τα σημεία αυτά είναι σειρές από n σύμβολα Ε και Α.

Η Κατανομή τ.μ. Χ σε ένα διωνυμικό πείραμα, δηλαδή η συνάρτηση p(x)=P(X=x) είναι

p(x)=P(X=x)=

η οποία λέγεται διωνυμική κατανομή $\binom{n}{x} p^x q^{n-x}, x = 0,1,2,...,n$

Τυχαία Μεταβλητή: Μια τ.μ. η οποία έχει την παραπάνω κατανομή λέγεται διωνυμική τ.μ. και συμβολίζεται Χ~Β(n,p).

Τα n,p λέγονται παράμετροι της διωνυμικής κατανομής

Μέση τιμή διωνυμικής κατανομής μ=Ε(x)=np

Διακύμανση διωνυμικής κατανομής σ^2=Var(x)=npq

Παράδειγμα 1

Έστω ότι οι απλές γεννήσεις παιδιών έχουν πιθανότητα 0.5 να φέρουν αγόρι. Υποθέτοντας ότι οι διάφορες γεννήσεις είναι ανεξάρτητες ως προς το φύλο να βρεθεί η πιθανότητα μια οικογένεια 6 παιδιών να έχει (1) ακριβώς 4 αγόρια, (2) τουλάχιστον 1 αγόρι και (3) το πολύ ένα αγόρι.

Λύση: Κάθε γέννηση θεωρείται ένα διωνυμικό πείραμα, με πιθανότητα επιτυχίας (αγόρι) p=0.5 και πιθανότητα αποτυχίας (κορίτσι) q=1-0.5=0.5

Έστω X τ.μ. που παριστάνει συμβατικά τον αριθμό των επιτυχιών (αγόρια)

(1)Η πιθανότητα να έχει η οικογένεια 4 αγόρια είναι

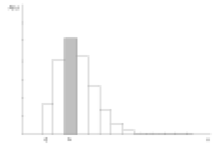
$$P(X=4)= \binom{6}{4}(0.5)^4(0.5)^{6-4} =$$

$$\frac{6!}{4!2!}(0.5)^6 = \frac{1*2*3*4*5*6}{1*2*3*4*1*2}(0.5)^6 =$$

0.2344

Ή χρησιμοποιώντας τους Στατιστικούς Πίνακες, σελ 3 ->

Πίνακας 2. Πιθανότητες Διωνυμικής Κατανομής



Τα στοιχεία του πίνακα εκφράζουν τις πιθανότητες $P(X=k)$, όπου η X ακολουθεί διωνυμική κατανομή με παραμέτρους n, p.

		p									
n	k	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
2	**0**	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.0950	0.1800	0.2550	0.3200	0.3750	0.4200	0.4550	0.4800	0.4950	0.5000
	2	0.0025	0.0100	0.0225	0.0400	0.0625	0.0900	0.1225	0.1600	0.2025	0.2500
3	**0**	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.1354	0.2430	0.3251	0.3840	0.4219	0.4410	0.4436	0.4320	0.4084	0.3750
	2	0.0071	0.0270	0.0574	0.0960	0.1406	0.1890	0.2389	0.2880	0.3341	0.3750
	3	0.0001	0.0010	0.0034	0.0080	0.0156	0.0270	0.0429	0.0640	0.0911	0.1250
4	**0**	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.1715	0.2916	0.3685	0.4096	0.4219	0.4116	0.3845	0.3456	0.2995	0.2500
	2	0.0135	0.0486	0.0975	0.1536	0.2109	0.2646	0.3105	0.3456	0.3675	0.3750
	3	0.0005	0.0036	0.0115	0.0256	0.0469	0.0756	0.1115	0.1536	0.2005	0.2500
	4	0.0000	0.0001	0.0005	0.0016	0.0039	0.0081	0.0150	0.0256	0.0410	0.0625
5	**0**	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313
	1	0.2036	0.3281	0.3915	0.4096	0.3955	0.3602	0.3124	0.2592	0.2059	0.1563
	2	0.0214	0.0729	0.1382	0.2048	0.2637	0.3087	0.3364	0.3456	0.3369	0.3125
	3	0.0011	0.0081	0.0244	0.0512	0.0879	0.1323	0.1811	0.2304	0.2757	0.3125
	4	0.0000	0.0005	0.0022	0.0064	0.0146	0.0284	0.0488	0.0768	0.1128	0.1563
	5	0.0000	0.0000	0.0001	0.0003	0.0010	0.0024	0.0053	0.0102	0.0185	0.0313
6	**0**	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
	1	0.2321	0.3543	0.3993	0.3932	0.3560	0.3025	0.2437	0.1866	0.1359	0.0938
	2	0.0305	0.0984	0.1762	0.2458	0.2966	0.3241	0.3280	0.3110	0.2780	0.2344
	3	0.0021	0.0146	0.0415	0.0819	0.1318	0.1852	0.2355	0.2765	0.3032	0.3125
	4	0.0001	0.0012	0.0055	0.0154	0.0330	0.0595	0.0951	0.1382	0.1861	0.2344
	5	0.0000	0.0001	0.0004	0.0015	0.0044	0.0102	0.0205	0.0369	0.0609	0.0938
	6	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0018	0.0041	0.0083	0.0156
7	**0**	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078

Παράδειγμα 1

(2) Η πιθανότητα να έχει τουλάχιστον ένα αγόρι

P (τουλάχιστον 1 αγόρι) = P (1 αγόρι και 5 κορίτσια)=
= P (2 αγόρια και 4 κορίτσια)=…= P(6γόρια)
= 1- P (κανένα αγόρι)=
= 1-P(X=0)
= 1-0,0156=0.9844

(3) Η πιθανότητα να έχει το πολύ ένα αγόρι
P (το πολύ 1 αγόρι) = P (1 αγόρι ή κανένα αγόρι) =
= P (1 αγόρι)+P(κανένα αγόρι)= 0,0938+0,0156=0,1094

Παράδειγμα 2

Έρευνες έδειξαν ότι το 40% των επισκεπτών μουσείων αγοράζουν και κάποιο αναμνηστικό αντικείμενο. Να υπολογιστεί η πιθανότητα ότι σε 12 τυχαία επιλεγμένους επισκέπτες ενός μουσείου

1.το πολύ 2 αγόρασαν κάποιο αναμνηστικό

2.τουλάχιστον 3 αγόρασαν κάποιο αναμνηστικό.

Έστω Χ ο αριθμός των επισκεπτών που αγοράζουν κάποιο αναμνηστικό.
Σύμφωνα με τα δεδομένα της άσκησης Χ ~ Διωνυμική Κατανομή με p = 0.40.

1. $P(x \leq 2/n=12,p=0.40) = P(x=0/n=12,p=0,40) + P(x=1/n=12,p=0,40) + P(x=2/n=12,p=0,40) = 0,0022+0,0174+0,00639 = 0.0835$

2. $P(x \geq 3/n=12, p=0.40) = 1 - P(x<3/n=12, p=0.40) =$
$1-[P(x=0/n=12, p=0.40)+ P(x=1/n=12, p=0.40)+P(x=2/n=12, p=0.40)]=$
$1-[0.0022 + 0.0174 + 0.0639] = 1-0.0835 = 0.9165 \cong 0.92$

Κατανομή Poisson[1]

Έστω μια ραδιενεργός ουσία η οποία με την πάροδο του χρόνου εκπέμπει σωματίδια ενός ορισμένου τύπου. Παρακολουθούμε την ουσία αυτή για ένα χρονικό διάστημα π.χ. μια ώρα ή ένα λεπτό και αναζητούμε τον αριθμό Χ των σωματιδίων που εκπέμπονται σ΄ αυτό το χρονικό διάστημα.

Χ: #των σωματιδίων είναι μια τ.μ. και ενδιαφερόμαστε να βρούμε την κατανομή της.

$X=0,1,2,3,\ldots,\infty$

Παραδοχές

(α) Οι συνθήκες του πειράματος παραμένουν σταθερές με την πάροδο του χρόνου

(β) Υποδιαιρούμε το χρονικό διάστημα σε μεγάλο αριθμό ίσων n υποδιαστημάτων, έτσι ώστε σε κάθε υποδιάστημα να εκπέμπεται <u>το πολύ ένα σωματίδιο</u> με θετική <u>πιθανότητα, p</u>.

(γ) Μπορούμε να πούμε ότι η εκπομπή των σωματιδίων μοιάζει με μια σειρά διωνυμικών κατανομών με επιτυχία Ε την εκπομπή ενός σωματιδίου και σταθερό p.

Άρα για μια κατά προσέγγιση τιμή της $P(X=x) = \binom{n}{x} p^x q^{n-x}$

Κατανομή Poisson

Έστω τώρα ότι ο αριθμός n→∞, p →0 και np=λ και σταθερή.

$$P(X=x)=\lim_{N\to\infty}\binom{n}{x}p^x q^{n-x}=\frac{\lambda^x}{x!}e^{-\lambda}$$

Η συνάρτηση $p_x(x)=P(X=x)=\dfrac{\lambda^x}{x!}e^{-\lambda}, x=0,1,2,\ldots,\lambda>0$

είναι κατανομή και ονομάζεται κατανομή Poisson με παράμετρο λ.

Η μέση τιμή της κατανομής Poisson είναι μ=E(X)=λ

Η διακύμανση της κατανομής Poisson Var(X)=λ

ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Συχνά η κατανομή Poisson αποτελεί προσέγγιση της διωνυμικής όταν $n \geq 20$ και $p\leq 0.05$, όταν $n \geq 100$ και $p\leq 0.01$ όταν ΔΗΛΑΔΗ λ=10.
2. Η κατανομή poisson χρησιμοποιείται σε περίπτωση ραδιενεργών διασπάσεων, τροχαίων ατυχημάτων, τηλεφωνικών κλήσεων που φτάνουν σε ένα κέντρο.

Παράδειγμα 3

Το τηλεφωνικό κέντρο ενός σταθμού πρώτων βοηθειών δέχεται τηλεφωνήματα έκτακτης ανάγκης που ακολουθούν την κατανομή Poisson με ρυθμό αφίξεων 2 τηλεφωνήματα ανά 30 λεπτά. Να βρεθεί η πιθανότητα το τηλεφωνικό κέντρο να δεχθεί τουλάχιστον 3 τέτοια τηλεφωνήματα από τις 12 το μεσημέρι μέχρι τις 2 μ.μ.

Έστω Χ ο αριθμός των τηλεφωνημάτων και ακολουθεί την κατανομή Poisson

Ως μονάδα χρόνου λαμβάνουνε την ώρα και εξετάζουμε το χρονικό διάστημα 12-2 μ.μ.άρα $\lambda=4*2=8$.

Η πιθανότητα που ζητάμε είναι $P(X\geq 3) = 1-P(X<3)$

$= 1-\{P(X=0)+P(X=1)+P(X=2)\}$

$$= 1 - \frac{8^0}{0!}e^{-8} - \frac{8^1}{1!}e^{-8} - \frac{8^2}{2!}e^{-8}$$

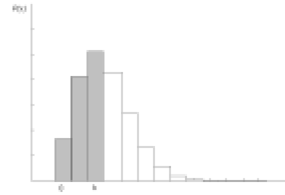
$$= 1 - e^{-8}\left(1+8+32\right)$$

$$= 1 - 41e^{-8} \cong 0.9863$$

Ή χρησιμοποιώντας τους Στατιστικούς Πίνακες, σελ 14 ->

Παράδειγμα 3

Πίνακας 5. Αθροιστικές Πιθανότητες Κατανομής Poisson
(συνέχεια)



						λ					
k	**7.5**	**8.0**	**8.5**	**9.0**	**9.5**	**10.0**	**12.0**	**15.0**	**18.0**	**20.0**	**25.0**
0	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
2	0.020	0.014	0.009	0.006	0.004	0.003	0.001	0.000	0.000	0.000	0.000
3	0.059	0.042	0.030	0.021	0.015	0.010	0.002	0.000	0.000	0.000	0.000
4	0.132	0.100	0.074	0.055	0.040	0.029	0.008	0.001	0.000	0.000	0.000
5	0.241	0.191	0.150	0.116	0.089	0.067	0.020	0.003	0.000	0.000	0.000
6	0.378	0.313	0.256	0.207	0.165	0.130	0.046	0.008	0.001	0.000	0.000
7	0.525	0.453	0.386	0.324	0.269	0.220	0.090	0.018	0.003	0.001	0.000
8	0.662	0.593	0.523	0.456	0.392	0.333	0.155	0.037	0.007	0.002	0.000
9	0.776	0.717	0.653	0.587	0.522	0.458	0.242	0.070	0.015	0.005	0.000
10	0.862	0.816	0.763	0.706	0.645	0.583	0.347	0.118	0.030	0.011	0.001
11	0.921	0.888	0.849	0.803	0.752	0.697	0.462	0.185	0.055	0.021	0.001
12	0.957	0.936	0.909	0.876	0.836	0.792	0.576	0.268	0.092	0.039	0.003
13	0.978	0.966	0.949	0.926	0.898	0.864	0.682	0.363	0.143	0.066	0.006
14	0.990	0.983	0.973	0.959	0.940	0.917	0.772	0.466	0.208	0.105	0.012
15	0.995	0.992	0.986	0.978	0.967	0.951	0.844	0.568	0.287	0.157	0.022
16	0.998	0.996	0.993	0.989	0.982	0.973	0.899	0.664	0.375	0.221	0.038

Παράδειγμα 4

Μία νέα γραμμή παραγωγής παρουσιάζει μέσο όρο 7 βλαβών ανά βδομάδα (5 εργασίμων ημερών). Να υπολογιστεί η πιθανότητα:

i) Σε μια τυχαία επιλεγμένη εβδομάδα να παρουσιαστούν το πολύ τρεις βλάβες

ii) Ότι σε ένα τυχαία επιλεγμένο διήμερο θα παρουσιαστούν τουλάχιστον 2 βλάβες

Έστω Χ ο αριθμός βλαβών ανά εβδομάδα
Σύμφωνα με τα δεδομένα της άσκησης Χ ~ Κατανομή Poisson με $\lambda = 7$

i. $P(x \leq 3/ \lambda=7) = P(x=0/ \lambda=7)+ P(x=1/ \lambda=7)+ P(x=2/ \lambda=7)+ P(x=3/ \lambda=7)=$
$0.0009 + 0.0064 + 0.0223 + 0.0521 = 0.0817 \cong 0.08$

ii. $\lambda = 7$ βλάβες ανά εβδομάδα 5 ημερών $\rightarrow \lambda_1 = \frac{7}{5} = 1.4$ βλάβες ανά ημέρα
$\rightarrow \lambda_2 = 1.4 * 2 = 2.8$ βλάβες ανά διήμερο

Έστω Υ ο αριθμός βλαβών ανά διήμερο
Σύμφωνα με τα δεδομένα της άσκησης Υ ~ Κατανομή Poisson με $\lambda_2 = 2.8$

$P(x \geq 2/ \lambda_1=2.8) = 1- P(x<2/ \lambda_1=2.8) =$
$1-[P(x=0/ \lambda_1=2.8) + P(x=1/ \lambda_1=2.8)] =$
$1-[0.0608+0.1703] = 1-0.2311 = 0.7689 \cong 0.77$

Κανονική Κατανομή [1]

Διακρίνουμε δύο περιπτώσεις κανονικής κατανομής

- **Τυπική (ή τυποποιημένη) κανονική** κατανομή με πυκνότητα

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < +\infty$$

X~N(0,1) και συμβολίζεται με Z

  – Μέση τιμή E(X)=0
  – Διακύμανση Var(X)=1

- **Γενική κανονική κατανομή** με πυκνότητα

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty, \quad -\infty < \mu < +\infty, \quad \sigma > 0$$

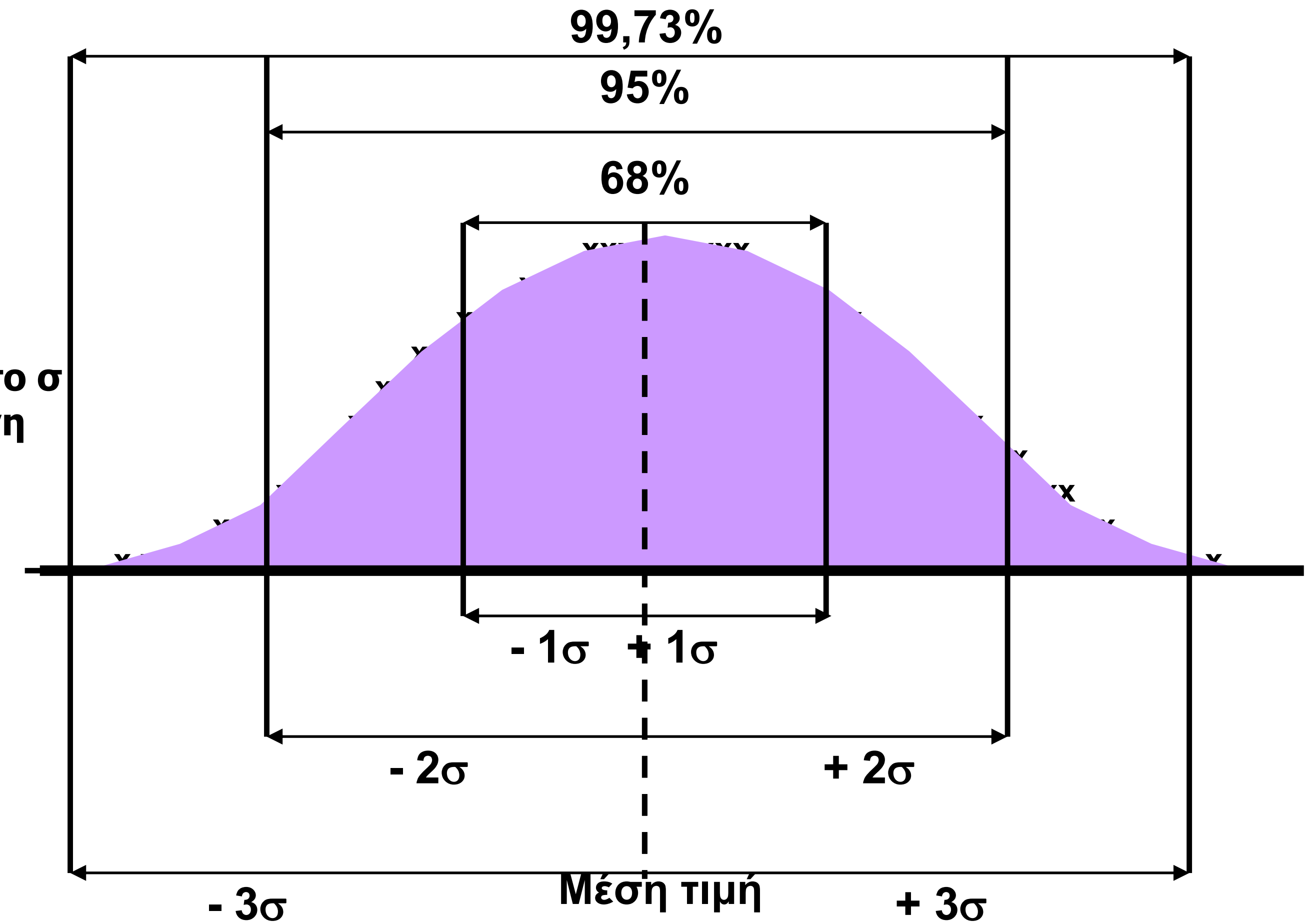
  – Μέση τιμή E(X)=μ
  – Διακύμανση Var(X)=σ^2

X~N(μ, σ^2)

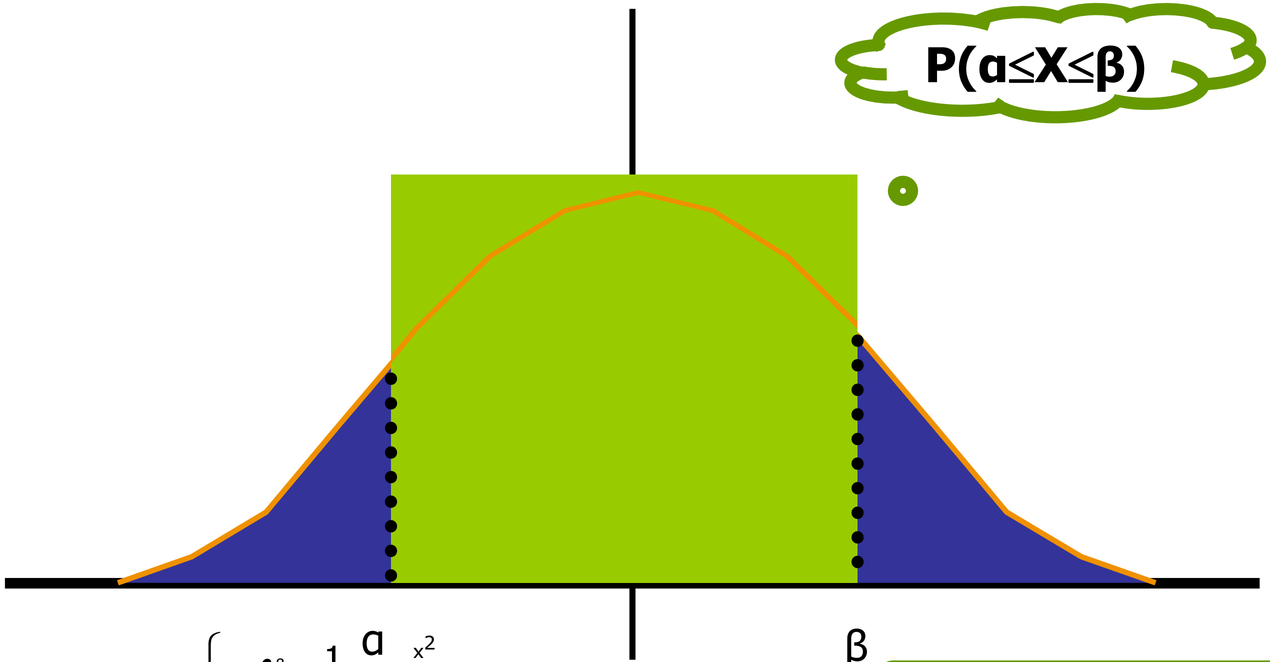
Προκύπτει

Κανονική κατανομή [2]

Οσο μεγαλύτερο το σ τόσο πιο απλωμένη είναι η f(x)



Κανονική κατανομή [3]



$$P(\alpha \le X \le \beta) = \int_{\alpha}^{\beta} f(x)dx = \begin{cases} \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx & X \sim N(0,1) \\ \int_{\alpha}^{\beta} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx & X \sim N(\mu,\sigma^2) \end{cases}$$

όταν η κατανομή δεν είναι τυπική την μετατρέπουμε

Η πιθανότητα υπολογίζεται από τους πίνακες της τυπικής κατανομής

Κανονική κατανομή [4]

Αν X~N(μ,σ²) τότε $Z = \dfrac{X-\mu}{\sigma}$ ~N(0,1)

Έτσι αν έχω να υπολογίσω την πιθανότητα στο διάστημα $\alpha \le X \le \beta$ όπου X~N(μ,σ²)

$$P(\alpha \le X \le \beta) = P\left(\frac{\alpha-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\beta-\mu}{\sigma}\right) = P(\alpha(\le X \le \beta^*)$$

οπου Z ~N(0,1)

Και η πιθανότητα υπολογίζεται πάλι από τους πίνακες.
Οι πίνακες δίνουν εμβαδόν

$$P(-\infty \le Z \le z_0) = \int_{-\infty}^{z_0} \frac{1}{\sqrt{2\pi}}^{e-\frac{z^2}{2}} dz \text{ ή}$$

$$P(z_0 \le Z \le \infty) = \int_{z_0}^{\infty} \frac{1}{\sqrt{2\pi}}^{e-\frac{z^2}{2}} dz, \quad z_0 \text{ πραγματικος αριθμός}$$

Λόγω της συμμετρίας θα χρησιμοποιήσουμε τους πίνακες που μας δίνουν εμβαδόν από 0 μέχρι z_0.

Παράδειγμα 5

Η κατανομή των μετρήσεων χοληστερίνης σε ένα πληθυσμό ακολουθούν την κανονική κατανομή με μέση τιμή 225 και τυπική απόκλιση 75. Να υπολογιστεί
(α) το ποσοστό των ατόμων με μετρήσεις χοληστερίνης μεταξύ 160 και 310

$$\mathrm{P}(160 < X < 310) = P\left(\frac{160\text{-}225}{75} \leq \frac{\mathrm{X} - \mu}{\sigma} \leq \frac{310\text{-}225}{75}\right) =$$
$$= \mathrm{P}(\text{-}0.87 < Z < 1.13)$$
$$= \mathrm{P}(Z < 1.13) - \mathrm{P}(\text{-}0.87)$$
$$= \ 0{,}8708 - 0.1922 = 0.6786$$

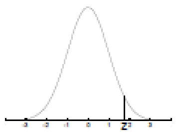
Χρησιμοποιώντας τους Στατιστικούς Πίνακες, σελ 15, 16 ->

Παράδειγμα 5

κάτω από την καμπύλη της τυποποιημένης κανονικής κατανομής αριστερά από το z .

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867

(συνέχεια)



$\Phi(z) = P(Z \leq z)$

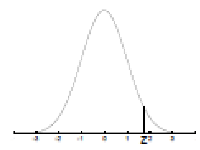
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	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

Παράδειγμα 5

(β) την τιμή της χοληστερίνης x0 τέτοια ώστε το 80% του πληθυσμού να έχει μικρότερη μέτρηση

$$P(X \le x_0) = 0.80 P\left(\frac{X-\mu}{\sigma} \le \frac{x_0-225}{75}\right) = 0.80 \Leftrightarrow P\left(Z \le \frac{x_0-225}{75}\right) = 0.80 \Leftrightarrow \textit{απο τους πινακες } z_0 = \frac{x_0-225}{75} = 0.84 \textit{Αρα } x_0 = 288$$

Πίνακας 6. Τυποποιημένη Κανονική Κατανομή
(συνέχεια)



$\Phi(z) = P(Z \le z)$

	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	0.7[illegible]	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

Παράδειγμα 6

Το βάρος των νεογέννητων παιδιών στη χώρα μας ακολουθεί την κανονική κατανομή με μέσο όρο 3 κιλά και 400 γραμμάρια για τα αγόρια και 3 κιλά και 350 γραμμάρια για τα κορίτσια και τυπική απόκλιση 0.4 κιλά για τα αγόρια και 0.350 για τα κορίτσια. Να υπολογισθούν οι πιθανότητες
(α) ένα νεογέννητο αγόρι να έχει βάρος μεγαλύτερο από 4 κιλά και
(β) ένα νεογέννητο κορίτσι να έχει βάρος μεγαλύτερο από 3 κιλά και μικρότερο από 3 κιλά και 700 γραμμάρια

ΛΥΣΗ:
Έστω Χ το βάρος του νεογέννητου αγοριού και Υ το βάρος του νεογέννητου κοριτσιού. Τότε

$$P(X > 4.000) = P\left(\frac{(X\text{-}3.400)}{400} > \frac{4.000\text{-}3.400}{400}\right) =$$

$$P(Z>1.5) = 1 - \Phi(1.5) = 1 - 0.9332 = 0.0668$$

$$P(3.000 < Y < 3700) = P\left(\frac{3.000\text{-}3.350}{350} < \frac{Y\text{-}3.350}{350} < \frac{3.700\text{-}3.350}{350}\right) =$$

$$P(-1<Z<1) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0.6826$$

Κατανομή Γάμμα

Έστω X συνεχής τυχαία μεταβλητή. Λέμε ότι η X ακολουθεί την **κατανομή γάμμα** με παραμέτρους α και β (α, β > 0) αν η συνάρτηση πυκνότητας πιθανότητας δίνεται από τον τύπο :

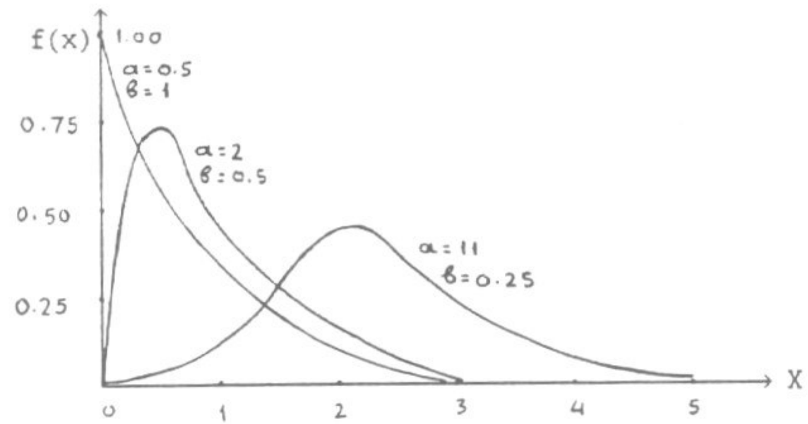
$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & \forall\, x > 0 \\ 0 & \forall\, x \leq 0 \end{cases}$$

όπου Γ(α) είναι η **συνάρτηση γάμμα** που εκφράζεται από το ολοκλήρωμα:

$$\Gamma(\alpha) = \int_{0}^{+\infty} x^{\alpha-1} e^{-x}\, dx.$$

Ισχύουν: $E(X) = \alpha\beta$

$V(X) = \alpha\beta^{2}$



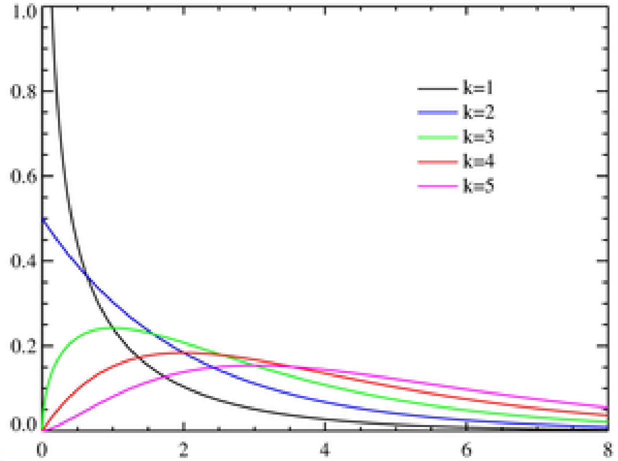
X^2-Κατανομή [1]

Έστω X μία συνεχής τυχαία μεταβλητή ή οποία ακολουθεί την κατανομή Γάμα με παραμέτρους $\alpha = \nu/2$, $\nu = 1,2,\ldots$ και $\beta = 2$.

Στην περίπτωση αυτή λέμε ότι η τυχαία μεταβλητή X ακολουθεί την κατανομή x^2 με ν βαθμούς ελευθερίας, συμβολίζεται δε $X \sim x^2_\nu$.

Το ν, δηλαδή οι βαθμοί ελευθερίας, είναι η μοναδική παράμετρος τής κατανομής.
Για την κατανομή x^2 αποδεικνύεται ότι ισχύουν οι σχέσεις:

$$E(X) = \nu \text{ και } \qquad V(X) = 2\nu$$



- Η γραφική παράσταση είναι λοξή και όσο οι
- βαθμοί ελευθερίας αυξάνονται ($\rightarrow\infty$) η x^2_ν
- κατανομή προσεγγίζει την κανονική

- Για τιμές του $\nu>30$ η x^2_ν προσεγγίζεται
- από την κανονική κατανομή

X^2-Κατανομή [2]

Λόγω της σπουδαιότητας που έχει η κατανομή x^2 στις εφαρμογές υπάρχουν έτοιμοι Πίνακες οι οποίοι για δοσμένη πιθανότητα δίνουν τα

(1-α) ποσοστιαία σημεία της κατανομής x^2 με ν βαθμούς ελευθερίας, δηλαδή τις τιμές για τις οποίες $\mathbf{P(\chi_v^2 \leq \chi_{a,v}^2)=1-a}$

Κατανομή Student (t) [1]

Η τυχαία κατανομή $X = \frac{Z}{\sqrt{Y/\nu}}$ όπου Z και Y είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομές την τυποποιημένη κανονική N(0,1) και την αντίστοιχα λέμε ότι X_{ν}^{2} ακολουθεί την κατανομή t του Student με ν βαθμούς ελευθερίας συμβολίζεται δε X ~ t_{ν}.

Το ν είναι η μοναδική παράμετρος της κατανομής.

Για την κατανομή t του Student αποδεικνύεται ότι ισχύουν:

$$E(X) = 0$$

και

$$V(X) = \frac{\nu}{\nu - 2}$$

Όσο οι βαθμοί ελευθερίας αυξάνονται η t_{ν} κατανομή προσεγγίζει την τυπική κανονική

Κατανομή Student (t) [2]

- Τα αντίστροφα (από πάνω) εκατοστιαία σημεία $t_{\alpha,v}$ της κατανομής t_v δίνονται από τους στατιστικούς πίνακες για τις διάφορες τιμές των βαθμών ελευθερίας ν.

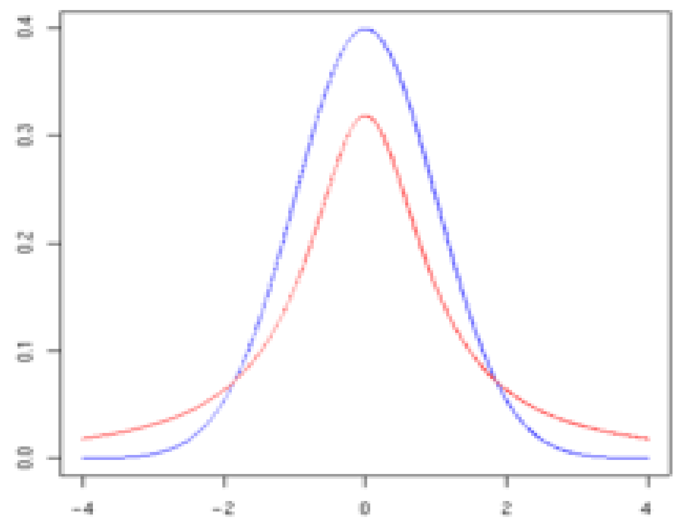
  Τα οποία ορίζονται από τη σχέση

$$\mathbf{P}(t_v \geq t_{\boldsymbol{\alpha},v})\mathbf{=\alpha}$$

Λόγω συμμετρίας έχουμε

$$P(t_v \leq - t_{\alpha,v}) = P(t_v \geq t_{\alpha,v}) = \alpha$$

$$t_{1-\alpha, v} \leq \mathbf{-}\, t_{\boldsymbol{\alpha},v}$$



Κατανομή *F* [1]

- **Ορισμός:** Έστω $X_1 \sim \chi_{v1}^2$, $X_2 \sim \chi_{v2}^2$ X_1, X_2 είναι ανεξάρτητες τ.μ. Η κατανομή της τ.μ.

$$Z = \frac{(X_1/v_1)}{(X_2/v_2)} \sim F_{v1,v2}$$

είναι η $F_{v1,v2}$ (F με v_1, v_2 βαθμούς ελευθερίας)

- Η κατανομή F είναι ασύμμετρη
- Η μέση τιμή και η διακύμανση της F δίνονται από τον τύπο

$$E\left(F_{v1,v2}\right) = \frac{\nu_2}{\nu_1 - 2}, \nu_1 > 2$$

$$Var\left(F_{v1,v2}\right) = \frac{2\nu^2{}_2\left(\nu_1 + \nu_2 - 2\right)}{\nu_1\left(\nu_2 - 2\right)^2\left(\nu_2 - 4\right)}, \nu_2 > 4$$

Κατανομή *F* [2]

- Λόγω της σπουδαιότητας που έχει η κατανομή F στη στατιστική συμπερασματολογία υπάρχουν έτοιμοι Πίνακες οι οποίοι δίνουν για δοσμένη πιθανότητα τα (1-α) ποσοστιαία σημεία της κατανομής F με ν1 και ν2 βαθμούς ελευθερίας,

$$P(F \leq F_{1-a\ v1,v2}) = 1-a$$

- Είναι προφανές ότι από τον Πίνακα προκύπτουν μόνο τα ανώτερα σημεία της κατανομής F. Τα κατώτερα ποσοστιαία σημεία προσδιορίζονται από τη σχέση

$$F_{1-a,v1,v2} = \frac{1}{F_{a,v1,v2}}$$