Κατανομές συνεχών τυχαίων μεταβλητών

Καθηγητής Χρήστος Γ. Μασούρος

Συνεχείς τυχαίες μεταβλητές

Οι κατανομές που χρησιμοποιούνται περισσότερο είναι:

- Η ομοιόμορφη κατανομή
- Η εκθετική κατανομή
- Η Κανονική κατανομή
- Η χ^2 κατανομή
- Η κατανομή t του Student
- Η F κατανομή

ομοιόμορφη κατανομή

- Η Συνάρτηση Πυκνότητας Πιθανότητας της ομοιόμορφης κατανομής δίνεται από τον τύπο:

$$f(x)=\begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

ομοιόμορφη κατανομή

- Η Συνάρτηση Κατανομής δίνεται από τον τύπο:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}$$

ομοιόμορφη κατανομή

- Η **μέση τιμή** και η **διακύμανση** μιας συνεχούς ομοιόμορφης κατανομής στο διάστημα [α,β] δίνεται από τους τύπους:

$$E(x)=\frac{a+b}{2}, \qquad V(x)=\frac{(b-a)^2}{2}$$

Παράδειγμα

Τα μαθήματα στο Πανεπιστήμιο αρχίζουν στις 9:15. Ένας φοιτητής που χρησιμοποιεί την αστική συγκοινωνία για να μεταβεί στο Πανεπιστήμιο ξεκινάει από την αφετηρία στις 8:00 κάθε πρωί. Ας υποθέσουμε ότι η διάρκεια του δρομολογίου μέχρι τον σταθμό του Πανεπιστημίου είναι μια τυχαία μεταβλητή που ακολουθεί ομοιόμορφη κατανομή στο διάστημα [58,63] (οι αριθμοί αφορούν min). Αν ο φοιτητής χρειάζεται επιπλέον 15min για να περπατήσει από τον σταθμό αποβίβασης μέχρι το αμφιθέατρο

Παράδειγμα

α) ποια η πιθανότητα να φτάσει στο αμφιθέατρο μετά την έναρξη του μαθήματος;

Αν συμβολίσουμε με Χ την διάρκεια του δρομολογίου, η τυχαία μεταβλητή Χ που ακολουθεί ομοιόμορφη κατανομή στο διάστημα [58,63] θα έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x)=\begin{cases}\dfrac{1}{63-58}, & 58\le x\le 63\\ 0, & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}=\begin{cases}\dfrac{1}{5}, & 58\le x\le 63\\ 0, & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}$$

Παράδειγμα

και συνάρτηση κατανομής

$$F(x)=\begin{cases}0, & x<58\\ \dfrac{x-a}{b-a}, & 58\leq x<63\\ 1, & x\geq 63\end{cases}$$

Παράδειγμα

Μετά την έναρξη του μαθήματος ο φοιτητής θα φτάσει στο αμφιθέατρο αν το δρομολόγιο διαρκέσει πάνω από 60 λεπτά. Ζητάμε συνεπώς την πιθανότητα

$$P(X>60) = 1-P(X\leq 60) = 1-F(60)$$

Όμως

$$F(60)=\frac{60-58}{5}=\frac{2}{5}$$

Άρα

$$P(x>60)=1-\frac{2}{5}=\frac{3}{5}=60\%$$

Παράδειγμα

β) ποια η πιθανότητα να φτάσει στο αμφιθέατρο τουλάχιστον 1 min πριν την έναρξη του μαθήματος;

Τουλάχιστον 1 min πριν την έναρξη του μαθήματος ο φοιτητής θα φτάσει στο αμφιθέατρο αν το δρομολόγιο διαρκέσει λιγότερο από 59 λεπτά. Ζητάμε συνεπώς την πιθανότητα

$$P\left(X \leq 59\right) = F\left(59\right) = \frac{59-58}{5} = \frac{1}{5} = 20\%$$

Παράδειγμα

γ) ποιος είναι ο αναμενόμενος χρόνος άφιξης του φοιτητή στο αμφιθέατρο;

Ο μέσος χρόνος του δρομολογίου είναι

$$E(X) = \frac{58+63}{2} = 60,5\,\text{min}$$

οπότε η αναμενόμενη στιγμή άφιξης του φοιτητή στο αμφιθέατρο θα είναι

$$60,5 + 15 = 75,5 \text{ min}$$

μετά τις 8:00, δηλαδή ο φοιτητής θα φτάνει στο μάθημα με καθυστέρηση μισού λεπτού.

Εκθετική κατανομή

- Η Συνάρτηση Πυκνότητας Πιθανότητας της εκθετικής κατανομής δίνεται από τον τύπο:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Εκθετική κατανομή

- Η Συνάρτηση Κατανομής δίνεται από τον τύπο:

$$F(t)=\begin{cases} 1-e^{-\lambda t}, & t\geq 0 \\ 0, & t<0 \end{cases}$$

Εκθετική κατανομή

- Η **μέση τιμή** και η **διακύμανση** μιας εκθετικής κατανομής στο διάστημα δίνεται από τους τύπους:

$$\mu = E(x) = \frac{1}{\lambda}, \qquad \sigma^2 = V(x) = \frac{1}{\lambda^2}$$

Παράδειγμα

Ο χρόνος ζωής Χ μιας συσκευής ακολουθεί την εκθετική κατανομή με παράμετρο λ>0.

(α) Να υπολογισθεί η πιθανότητα ο χρόνος ζωής της συσκευής να υπερβεί το μέσο χρόνο ζωής της.

Αν συμβολίσουμε με F τη συνάρτηση κατανομής της τυχαίας μεταβλητής Χ θα έχουμε

$$F(t) = P(X \leq t) = 1 - e^{-\lambda t}, \quad t \geq 0$$

Ενώ για τον μέσο χρόνο ζωής γνωρίζουμε ότι E(X)=1/λ. Επομένως

$$P(X > E(X)) = P\left(X > \frac{1}{\lambda}\right) = 1 - P\left(X \leq \frac{1}{\lambda}\right) = 1 - F\left(\frac{1}{\lambda}\right) =$$

$$= 1 - \left(1 - e^{-\lambda \frac{1}{\lambda}}\right) = 1 - \left(1 - e^{-1}\right) = e^{-1} \cong 0,3679$$

Παράδειγμα

(β) Να υπολογισθεί η πιθανότητα ο χρόνος ζωής της συσκευής να υπερβεί το διπλάσιο του μέσου χρόνου ζωής της.

$$P\left(X>2E\left(X\right)\right)=P\left(X>\frac{2}{\lambda}\right)=1-P\left(X\le\frac{2}{\lambda}\right)=1-F\left(\frac{2}{\lambda}\right)=$$

$$=1-\left(1-e^{-\lambda\frac{2}{\lambda}}\right)=1-\left(1-e^{-2}\right)=e^{-2}\cong 0{,}1353$$

Παράδειγμα

(γ) Να βρεθεί ποιος είναι ο χρόνος τον οποίο υπερβαίνει το 95% των παραγομένων συσκευών.

Αναζητούμε έναν αριθμό $x_o>0$ για τον οποίο ισχύει:

$$P(X>x_o) > 0,95$$

ή ισοδύναμα

$$1-F\left(x_o\right)=0,95 \Leftrightarrow 1-\left(1-e^{-\lambda x_o}\right)=0,95 \Leftrightarrow e^{-\lambda x_o}=0,95 \Leftrightarrow$$

$$\Leftrightarrow -\lambda x_o=\ln 0,95 \Leftrightarrow x_o=-\frac{1}{\lambda}\ln 0,95 \cong \frac{0,0513}{\lambda}$$

Κανονική κατανομή

- Η Συνάρτηση Πυκνότητας Πιθανότητας της κανονικής κατανομής δίνεται από τον τύπο:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad , \qquad \gamma\iota\alpha - \infty < x < +\infty,$$

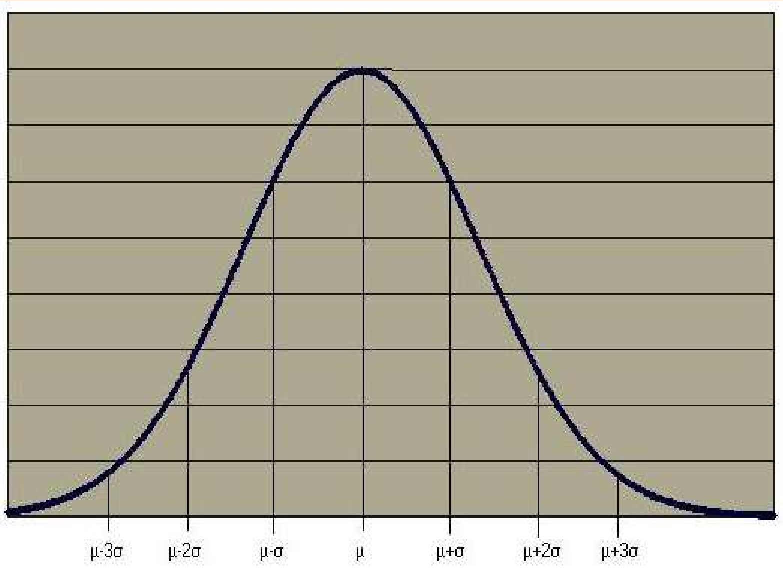
- Η Συνάρτηση Κατανομής δίνεται από τον τύπο:

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = P(X < x)$$

Κανονική κατανομή

- Αν γνωρίζουμε την μέση τιμή και την διασπορά της Κανονικής κατανομής, γνωρίζουμε πλήρως και την ίδια την κατανομή. Αν γνωρίζουμε δηλαδή τις δύο παραμέτρους της Κανονικής κατανομής γνωρίζουμε απολύτως την κατανομή.
- Συμβολίζουμε με $N(\mu, \sigma^2)$ την Κανονική κατανομή και γράφουμε $X \sim N(\mu, \sigma^2)$ για να δηλώσουμε ότι η τυχαία μεταβλητή Χ ακολουθεί την *Κανονική κατανομή* ή κατανέμεται σύμφωνα με την *Κανονική κατανομή*.

Ιδιότητες της κανονικής κατανομής

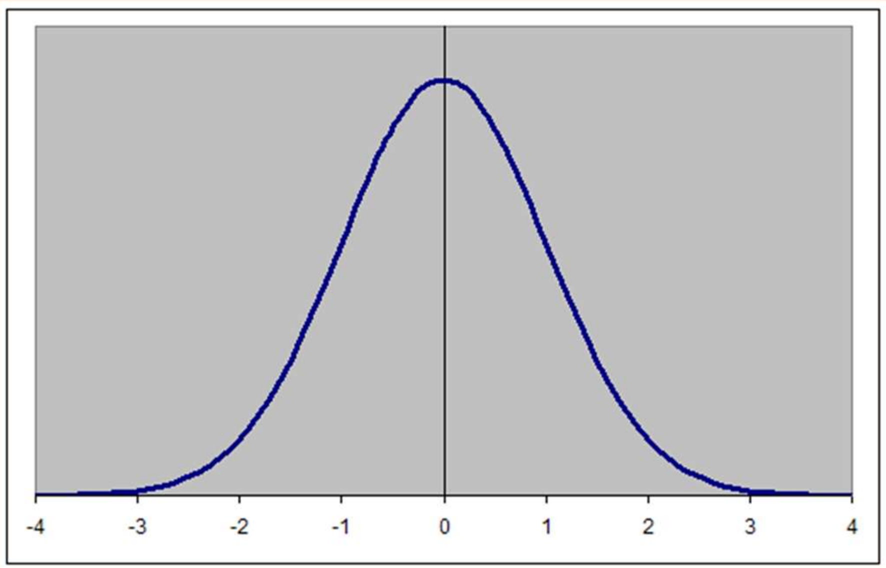


- Η συνάρτηση πυκνότητας πιθανότητας έχει συμμετρική γραφική παράσταση περί την μέση τιμή.
- Η μέση τιμή, η διάμεσος και η επικρατούσα τιμή ταυτίζονται.
- Το σχήμα της κατανομής είναι σαν καμπάνα (κωδωνοειδές).
- Η συμμετρία και η ταύτιση μέσης τιμής, διαμέσου και επικρατούσας τιμής δεν είναι ιδιότητα μόνον της Κανονικής κατανομής. Ωστόσο οι εξής τρεις ιδιότητες είναι:
- στο διάστημα (μ-σ, μ+σ) ανήκει περίπου το 68% των τιμών της Χ
- στο διάστημα (μ-2σ, μ+2σ) ανήκει περίπου το 95% των τιμών της Χ
- στο διάστημα (μ-3σ, μ+3σ) ανήκει περίπου το 99% των τιμών της Χ

Τυπική Κανονική Κατανομή

- Η κανονική κατανομή με παραμέτρους **$\mu=0$** και **$\sigma^2=1$** ονομάζεται *Τυπική Κανονική Κατανομή* και είναι ιδιαίτερα σημαντική, και συμβολίζεται με N(0,1). Η τυχαία μεταβλητή Ζ που ακολουθεί την Τυπική κανονική κατανομή ονομάζεται Τυπική τυχαία μεταβλητή.
- Ένας λόγος, για τον οποίο είναι σημαντική η τυπική κανονική κατανομή, είναι το ότι κάθε κανονική κατανομή ανάγεται στην τυπική κανονική κατανομή μέσω του *Μετασχηματισμού Τυποποίησης* ή απλά της *Τυποποίησης*.
- Με το μετασχηματισμό αυτόν αφαιρούμε από κάθε τιμή της αρχικής τυχαίας μεταβλητής Χ (που κατανέμεται κανονικά με παραμέτρους μ και σ^2) τη μέση τιμή μ και διαιρούμε με την τυπική απόκλιση σ.

Τυπική κανονική κατανομή



- Το αποτέλεσμα της τυποποίησης είναι:
$$Z = \frac{X - \mu}{\sigma}$$
- Για την τυπική κανονική κατανομή έχουμε ότι:
- στο διάστημα (-1, 1) ανήκει περίπου το 68% των τιμών της Z
- στο διάστημα (-2, 2) ανήκει περίπου το 95% των τιμών της Z
- στο διάστημα (-3, 3) ανήκει περίπου το 99% των τιμών της Z

Παράδειγμα
Πώς γίνεται η τυποποίηση

- Έστω X~N(25,9). Δηλαδή $\mu=25$ και $\sigma^2=9$, άρα $\sigma=3$. Επομένως έχουμε:

$$Z = \frac{X-\mu}{\sigma} = \frac{X-25}{3}$$

Συναρτήσεις πιθανότητας της τυπικής κανονικής κατανομής

- Συνάρτηση πυκνότητας πιθανότητας της Τυπικής μεταβλητής

- $\mathrm{f(z)} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}} \quad \text{για} - \infty < \mathrm{z} < +\infty$

- Συνάρτηση κατανομής της τυπικής μεταβλητής

- $\mathrm{F_Z(z)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{1}{2}t^2\right) dt = \mathrm{P(Z < z)}$

Ιδιότητες

- Όταν X~N(μ, σ²) τότε $F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right)$
- δηλαδή η αθροιστική πιθανότητα της X ισούται με την αθροιστική πιθανότητα της τυπικής μεταβλητής της X στη οποία μετασχηματίστηκε μέσω του μετασχηματισμού τυποποίησης. Και επειδή για οποιαδήποτε τυχαία μεταβλητή X ισχύει:
- $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$
- $P(x_1 \leq X \leq x_2) = F_Z\left(\frac{x_2-\mu}{\sigma}\right) - F_Z\left(\frac{x_1-\mu}{\sigma}\right)$

Ιδιότητες

- Από αυτό το σημείο και μετά ανατρέχουμε στους πίνακες της αθροιστικής τυπικής κανονικής κατανομής.
- Αν πρέπει να υπολογίσουμε την αθροιστική πιθανότητα για αρνητικές τιμές της Z έχουμε:

- $F_Z(-z^*) = P(Z \leq -z^*) = P(Z \geq z^*) = 1 - F_Z(z^*)$

- Το παραπάνω ισχύει λόγω της συμμετρίας της συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής. Πάντα βοηθάει να κάνουμε μια γραφική παράσταση ώστε να έχουμε εικόνα της συνάρτησης πυκνότητας πιθανότητας του σημείου z ή x και της πιθανότητας που αναζητούμε.

Εύρεση της πιθανότητας από τον πίνακα πιθανοτήτων της τυπικής κανονικής κατανομής

θέλουμε να υπολογίσουμε την πιθανότητα P(Z≤1,65)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5	0,504	0,508	0,512	0,516	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	0,591	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	0,648	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,67	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,695	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,719	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	0,758	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,791	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	0,834	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	0,877	0,879	0,881	0,883
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,898	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	0,937	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	0,975	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	0,983	0,9834	0,9838	0,9842	0,9846	0,985	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,989
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	0,992	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,994	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	0,996	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,997	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	0,998	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	0,999	0,999
3,1	0,999	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

Μερικές χρήσιμες συμβουλές

- Υπολογίζουμε πιθανότητες μόνο για διαστήματα και όχι για συγκεκριμένες τιμές της X. Η πιθανότητα $P(X=k)$ ισούται με μηδέν
- Στις συνεχείς τυχαίες μεταβλητές, δεν επηρεάζει τα αποτελέσματα αν χρησιμοποιούμε ανισότητες ή ανισοϊσότητες όταν καθορίζουμε τα διαστήματα για τα οποία θέλουμε να υπολογίσουμε την πιθανότητα
- Για να υπολογίσουμε μια πιθανότητα σχετικά με την τυχαία μεταβλητή που ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$ εργαζόμαστε ως εξής:
  - ✓ πρώτα κάνουμε τον μετασχηματισμό τυποποίησης-υπολογίζουμε δηλαδή τις z-τιμές,
  - ✓ μετά φροντίζουμε η φορά της ανισότητας να είναι προς τα αριστερά (δηλ $Z<z$), και
  - ✓ τέλος αν το z είναι αρνητικό κάνουμε το μετασχηματισμό $P(Z<z)=1-P(Z<-z)$ ώστε να έχουμε τη θετική τιμή $-z$
- Επίσης για την περίπτωση που θέλουμε να υπολογίσουμε την πιθανότητα $P(z_1<Z<z_2)$ χρησιμοποιούμε την ιδιότητα $P(z_1<Z<z_2)=P(Z<z_2)-P(Z<z_1)$.

Παράδειγμα

Αν η τυχαία μεταβλητή Χ ακολουθεί την κανονική κατανομή Ν(10,4) υπολογίστε τις πιθανότητες:

α) $P(X<10)$,

β) $P(X<12)$,

γ) $P(X<8)$,

δ) $P(7<X<12)$,

ε) $P(X>8)$,

στ) $P(X=12)$,

ζ) $P(X>13)$.

Παράδειγμα

- α) P(X<10)=0,5 αφού το 10 είναι η μέση τιμή και διάμεσος της X.

- β) $\mathrm{P}(\mathrm{X} < 12) = \mathrm{P}\left(\frac{\mathrm{X}-\mu}{\sigma} < \frac{12-\mu}{\sigma}\right) = \mathrm{P}\left(\mathrm{Z} < \frac{12-10}{2}\right) = \mathrm{P}(\mathrm{Z} < 1) =$

  $= \mathrm{F_Z}(1) = 0{,}8413$

- γ) $\mathrm{P}(\mathrm{X} < 8) = \mathrm{P}\left(\frac{\mathrm{X}-\mu}{\sigma} < \frac{8-\mu}{\sigma}\right) = \mathrm{P}\left(\mathrm{Z} < \frac{8-10}{2}\right) = \mathrm{P}(\mathrm{Z} < -1) =$

  $= 1 - \mathrm{P}(\mathrm{Z} < 1) = 1 - \mathrm{F_Z}(1) = 1 - 0{,}8413 = 0{,}1587$

Παράδειγμα

- δ) $P(7 < X < 12) = P\left(\frac{7-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{12-\mu}{\sigma}\right) =$

$$= P\left(\frac{7-10}{2} < \frac{X-\mu}{\sigma} < \frac{12-10}{2}\right) = P(-1{,}5 < Z < 1)$$

$$= P(Z < 1) - P(Z < -1{,}5) = P(Z < 1) - [1 - P(Z < 1{,}5)]$$

$$= P(Z < 1) - 1 + P(Z < 1{,}5) = F_Z(1) + F_Z(1{,}5) - 1$$

$$= 0{,}8413 + 0{,}9332 - 1 = 0{,}7745$$

Παράδειγμα

- ε) $P(X > 8) = P\left(\frac{X-\mu}{\sigma} > \frac{8-\mu}{\sigma}\right) = P\left(Z > \frac{8-10}{2}\right) = P(Z > -1) =$
  $P(Z < 1) = 0{,}8413$

- στ) $P(X = 12) = 0$

- ζ) $P(X > 13) = P\left(\frac{X-\mu}{\sigma} > \frac{13-\mu}{\sigma}\right) = P\left(Z > \frac{13-10}{2}\right) = P(Z > 1{,}5) =$
  $1 - P(Z \leq 1{,}5) = 1 - F_Z(1{,}5) = 1 - 0{,}9332 = 0{,}0668$

Παράδειγμα

Το ύψος των κατοίκων μιας πόλης με πληθυσμό 1000, ακολουθεί την κανονική κατανομή με μέση τιμή μ=170cm και τυπική απόκλιση σ=7cm. Υπολογίστε τον αριθμό των κατοίκων με ύψος

α. κάτω των 183cm

β. μεταξύ 165cm και 177cm

γ. κάτω των 160cm ή άνω των 180cm

<u>Λύση</u>

Έστω X τυχαία μεταβλητή η οποία εκφράζει το ύψος των κατοίκων της πόλης. Τότε $X \sim N(170, 7^2)$. Θα υπολογίσουμε αρχικά τις πιθανότητες:

α. $P(X<183)$

β. $P(165<X<177)$

γ. $P(X<160$ ή $X>180)$

Παράδειγμα

- α) $P(X < 183) = P\left(\frac{X-\mu}{\sigma} < \frac{183-\mu}{\sigma}\right) = P\left(Z < \frac{183-}{7}\right) = P(Z < 1{,}86) =$
$= F_Z\ (1{,}86) = 0{,}9686$

- β) $P(165\ <\ X\ <\ 177) = P(\ X\ <\ 177) - P(X\ \leq\ 165) =$

$$= P\left(Z < \frac{177-170}{7}\right) - P\left(Z < \frac{165-170}{7}\right) =$$

$$=P(Z < 1) - P(Z < -0{,}71) = P(Z < 1) - \big(1 - P(Z < 0{,}71)\big) =$$

$$= P(Z < 1) + P(Z < 0{,}71) - 1 = F_Z(1) + F_Z(0{,}71) - 1 =$$

$$= 0{,}8413 + 0{,}7611 - 1 = 0{,}6024$$

Παράδειγμα

$$\gamma)\ \ \mathrm{P}(\mathrm{X} < 160 \text{ ή } \mathrm{X} > 180)$$
$$= \mathrm{P}(\mathrm{X} < 160) + \mathrm{P}(\mathrm{X} > 180)$$
$$= \mathrm{P}(\mathrm{X} < 160) + 1 - \mathrm{P}(\mathrm{X} \leq 180)$$
$$= 1 + \ \mathrm{P}\left(\mathrm{Z} < \frac{160 - 170}{7}\right) - \mathrm{P}\left(\mathrm{Z} < \frac{180 - 170}{7}\right)$$
$$= 1 + \mathrm{P}(\mathrm{Z} < -1{,}43) - \mathrm{P}(\mathrm{Z} < 1{,}43)$$
$$= 1 + \big(1 - \mathrm{P}(\mathrm{Z} < 1{,}43)\big) - \mathrm{P}(\mathrm{Z} < 1{,}43)$$
$$= 2 - 2\mathrm{P}(\mathrm{Z} < 1{,}43) = 2 - 2 \cdot 0{,}9236 = 0{,}1528$$

Παράδειγμα

Επομένως:

α) κάτοικοι με ύψος κάτω των 183cm είναι:

$1.000 \times 0{,}9686 = 968{,}6 \approx 969$

β) κάτοικοι με ύψος μεταξύ 165cm και 177cm

είναι: $1.000 \times 0{,}6024 = 602{,}4 \approx 602$

γ) κάτω των 160cm ή άνω των 180cm είναι:

$1.000 \times 0{,}1528 = 152{,}8 \approx 153$

Παράδειγμα

i. Έστω X η τυχαία μεταβλητή που δηλώνει το ενοίκιο ανά τετραγωνικό μέτρο (σε ευρώ) καταστημάτων στην συγκεκριμένη οδό. Τότε $X \sim N(30,5^2)$.

$$P(20 \leq X \leq 35) = P\left(\frac{20-30}{5} \leq Z \leq \frac{35-30}{5}\right) =$$

$$= P(-2 \leq Z \leq 1) = \Phi(1) - \Phi(-2) =$$

$$= 0{,}8413 - 0{,}0228 = 0{,}8185.$$

Παράδειγμα

ii. Έστω x η ζητούμενη τιμή. Τότε:

$P(X>x) = 0,10 \Leftrightarrow P(Z > \frac{x-30}{5}) = 0,01 \Leftrightarrow$

$\Leftrightarrow 1 - P(Z < \frac{x-30}{5}) = 0,01 \Leftrightarrow$

$\Leftrightarrow P(Z < \frac{x-30}{5}) = 1-0,01 \Leftrightarrow$

$\Leftrightarrow P(Z < \frac{x-30}{5}) = 0,9$

Παράδειγμα

Από τους πίνακες της κανονικής κατανομής βλέπουμε ότι Φ(1,28)=0,8997≈0,90

	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

Παράδειγμα

Συνεπώς

$$\frac{x-30}{5} = 1{,}28 \Leftrightarrow x = 1{,}28 \cdot 5 + 30 \Leftrightarrow x = 36{,}4$$

Επομένως η τιμή ενοικίασης ανά τ.μ. που μόνο το 10% των καταστημάτων την ξεπερνούν είναι 36,4 €.

Η κατανομή χ^2

την κατανομή χ^2 εισήγαγε ο F. R. Helmert το 1876

Αν Z_1 , Z_2,…, Z_n είναι *n ανεξάρτητες τυποποιημένες κανονικές* τυχαίες μεταβλητές, δηλαδή, αν

$$Z_i \sim N\,(0, 1),\ i = 1,2,\ldots,n$$

τότε η κατανομή της τυχαίας μεταβλητής,

$$X = Z_1^2 + Z_2^2 + \ldots + Z_n^2$$

ονομάζεται ***κατανομή χι-τετράγωνο με n βαθμούς ελευθερίας*** και συμβολίζεται με χ_n^2.

Η κατανομή χ^2

Είναι προφανές ότι πρόκειται για οικογένεια κατανομών. Για κάθε τιμή του *n* παίρνουμε και μια άλλη *κατανομή χι-τετράγωνο*. Είναι επίσης προφανές ότι μια τυχαία μεταβλητή *X* που ακολουθεί μια χ_n^2 *κατανομή* δεν παίρνει αρνητικές τιμές.

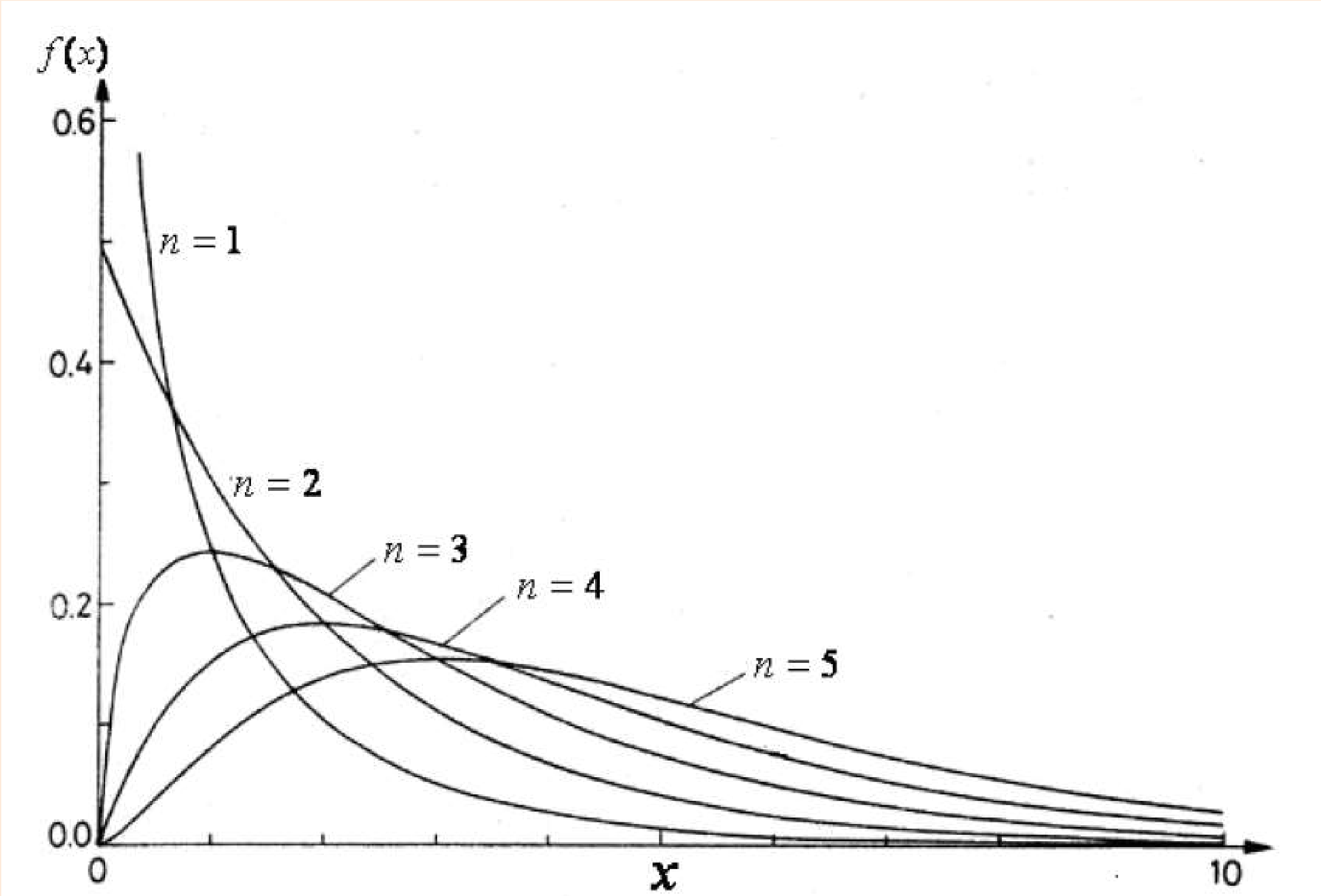
Στο σχήμα που ακολουθεί φαίνεται η γραφική παράσταση της *συνάρτησης πυκνότητας* της

$$X = Z_1^2 + Z_2^2 + \ldots + Z_n^2 \sim \boldsymbol{\chi_n^2}$$

για διάφορες τιμές του *n*.

Η κατανομή χ^2

Παρατηρείστε στο παρακάνω σχήμα ότι όσο το *n* αυξάνεται τόσο η γραφική παράσταση της *συνάρτησης πυκνότητας* της χ_n^2 γίνεται πιο συμμετρική. Για $n > 30$, προσεγγίζεται πολύ ικανοποιητικά από την *κανονική* $N(n, 2n)$.



Η κατανομή χ^2

Αποδεικνύεται ότι η μέση τιμή της χ_n^2 είναι ίση με **n** και η διασπορά της είναι ίση με **2n**.
Δηλαδή, αν $X \sim \chi_n^2$ τότε

$$E(X) = n$$

και

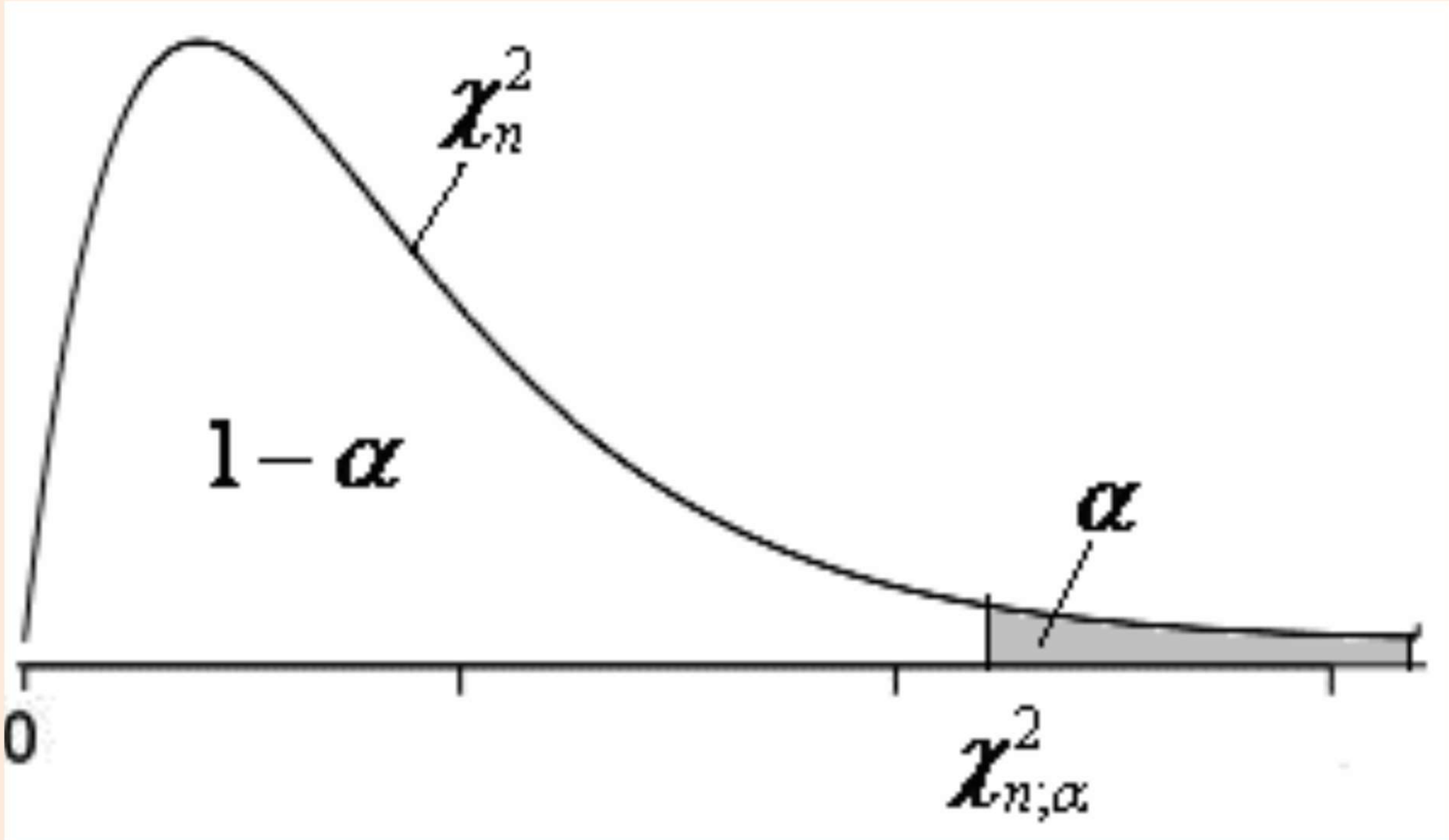
$$V(X) = 2n$$

Η κατανομή χ^2

Στη *Στατιστική Συμπερασματολογία* είναι χρήσιμα τα *άνω α -ποσοστιαία σημεία* της χ_n^2 τα οποία έχει επικρατήσει να συμβολίζονται με $\chi^2_{n;\alpha}$ ή με $\chi_n^2(\alpha)$. Όπως φαίνεται στο σχήμα που ακολουθεί, αν μια τυχαία μεταβλητή *X* ακολουθεί την κατανομή χ_n^2, τότε το $\chi^2_{n;\alpha}$ είναι εκείνη η τιμή της *X* για την οποία ισχύει

$P(X > \chi^2_{n;\alpha}) = \alpha$, ή ισοδύναμα, **$P(X \leq \chi^2_{n;\alpha}) = 1-\alpha$**

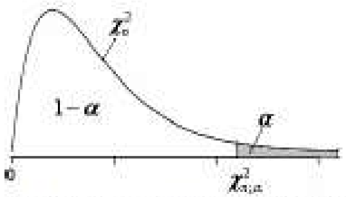
Η κατανομή χ^2



Τιμές $\chi^2_{n;\alpha}$ της κατανομής χ^2_n

Ο Πίνακας δίνει τα άνω α -ποσοστιαία σημεία της κατανομής χ^2 με n βαθμούς ελευθερίας

Αν $X \sim \chi^2_n$, ισχύει, $P(X > \chi^2_{n;\alpha}) = \alpha$.



n	α = 0.995	α = 0.99	α = 0.975	α = 0.95	α = 0.05	α = 0.025	α = 0.01	α = 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	24.4331	26.509	55.756	59.342	63.691	66.766
50	27.991	29.708	32.3574	34.764	67.505	71.420	76.154	79.490
60	35.535	37.485	40.4817	43.188	79.082	83.298	88.379	91.952
70	43.275	45.442	48.7576	51.739	90.531	95.023	100.425	104.215
80	51.172	53.540	57.1532	60.392	101.879	106.629	112.329	116.321
90	59.196	61.754	65.6466	69.126	113.145	118.136	124.116	128.299
100	67.328	70.065	74.2219	77.930	124.342	129.561	135.807	140.169

για παράδειγμα:

$\chi^2_{4;0,05}$ = 9,488

$\chi^2_{5;0,01}$ = 15,086

$\chi^2_{7;0,995}$ = 0,989

Η κατανομή t ή κατανομή Student

την κατανομή αυτή εισήγαγε ο W. L. Gosset το 1876
Είναι γνωστή ως κατανομή Student από το ψευδώνυμο
με το οποίο ο Gosset δημοσίευε τα άρθρα του.

Έστω *Z* μια τυχαία μεταβλητή η οποία ακολουθεί την *τυποποιημένη κανονική κατανομή*, δηλαδή

$$Z \sim N(0, 1)$$

και S_n μια τυχαία μεταβλητή *ανεξάρτητη* από την *Z* η οποία ακολουθεί την κατανομή χ_n^2, δηλαδή

$$S_n \sim \chi_n^2$$

Η κατανομή t ή κατανομή Student

Τότε, η κατανομή της τυχαίας μεταβλητής

$$\boldsymbol{T = \frac{Z}{\sqrt{\frac{S_n}{n}}}}$$

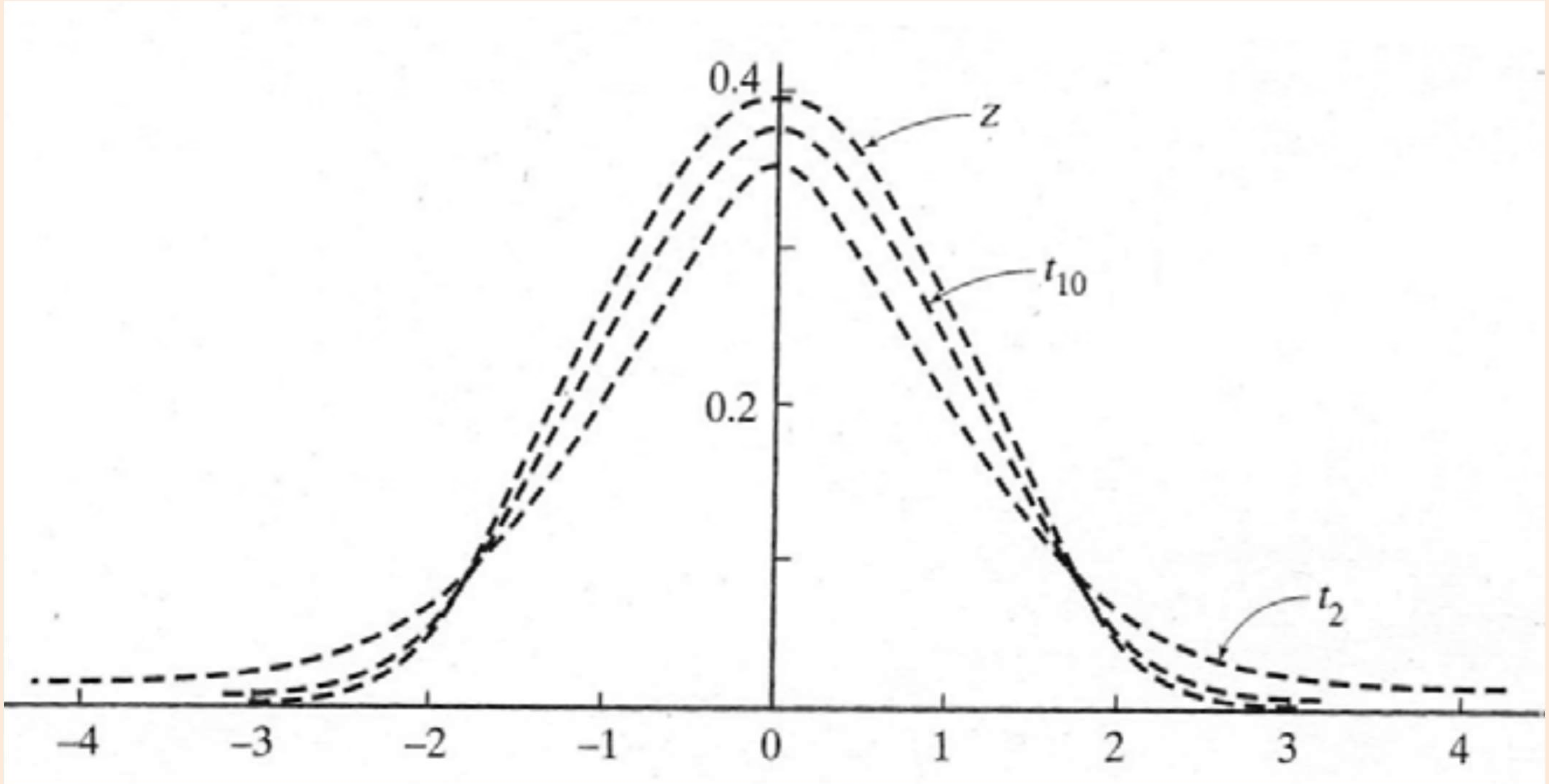
ονομάζεται ***κατανομή t ή κατανομή Student με n βαθμούς ελευθερίας*** και συμβολίζεται με t_n .

- Είναι προφανές ότι πρόκειται για οικογένεια κατανομών. Για κάθε τιμή του *n* παίρνουμε και μια άλλη *κατανομή t*.
- Είναι επίσης προφανές ότι μια τυχαία μεταβλητή *T* που ακολουθεί μια t_n *κατανομή* παίρνει τιμές από $-\infty$ έως $+\infty$ όπως η *κανονική κατανομή*.

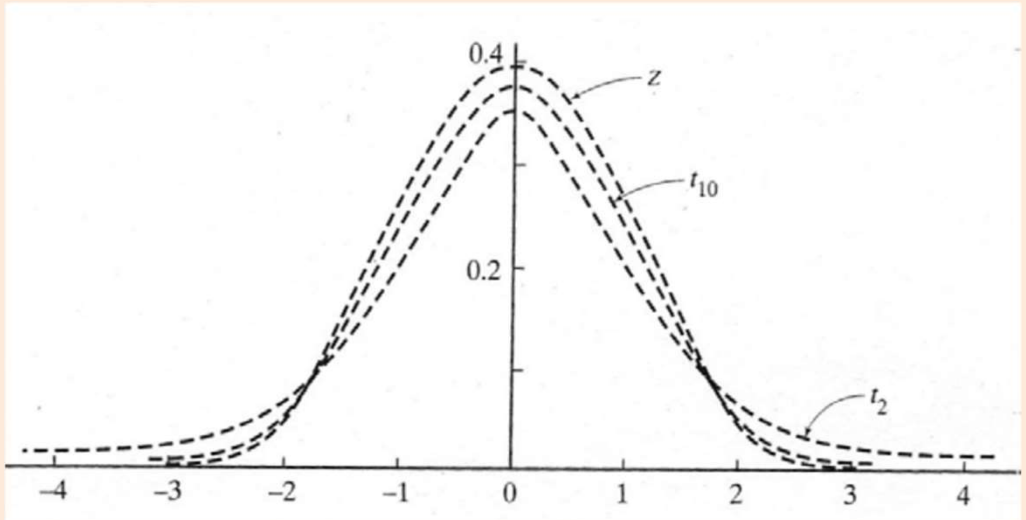
Η κατανομή t ή κατανομή Student

Στο παρακάνω σχήμα φαίνεται η γραφική παράσταση της *συνάρτησης πυκνότητας* της $T \sim t_n$ για $n = 2$ και για $n = 10$.

Επίσης φαίνεται η γραφική παράσταση της *συνάρτησης πυκνότητας* της $Z \sim N(0, 1)$.



Η κατανομή t ή κατανομή Student



Παρατηρείστε ότι η γραφική παράσταση της *συνάρτησης πυκνότητας* της $T \sim t_n$ έχει *κωδωνοειδή* μορφή και είναι συμμετρική ως προς τον κατακόρυφο άξονα στο 0, όπως η $Z \sim N(0, 1)$, όμως έχει πιο «παχιές» ουρές (είναι πιο πεπλατυσμένη).

Όσο το n αυξάνεται η t_n προσεγγίζεται ικανοποιητικά από την *κανονική* $N(0, n/(n-2))$.

Για $n > 30$, προσεγγίζεται πολύ καλά από την *τυποποιημένη κανονική* $N(0, 1)$.

Η κατανομή t ή κατανομή Student

Αποδεικνύεται ότι η μέση τιμή της t_n είναι ίση με **0** (για **n>1**) και η διασπορά της είναι ίση με **n/(n-2)** (για **n>2**). Δηλαδή, αν $T \sim t_n$ τότε

$$E(T) = 0 \text{ , για } n>1$$

και

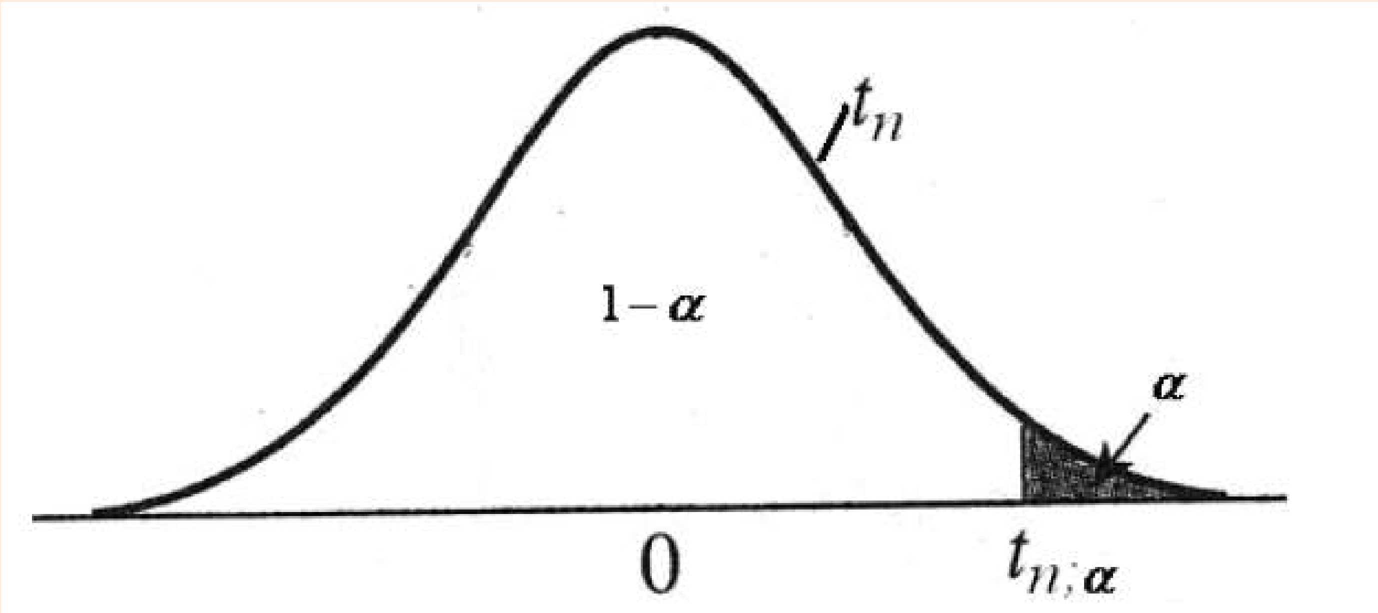
$$V(T) = n/(n-2) \text{ , για } n>2$$

Η κατανομή t ή κατανομή Student

Στη *Στατιστική Συμπερασματολογία* είναι χρήσιμα τα *άνω* α *-ποσοστιαία σημεία* της t_n τα οποία έχει επικρατήσει να συμβολίζονται με $t_{n;\alpha}$ ή με $t_n(\alpha)$. Όπως φαίνεται στο σχήμα που ακολουθεί, αν μια τυχαία μεταβλητή *T* ακολουθεί την κατανομή t_n, τότε το $t_{n;\alpha}$ είναι εκείνη η τιμή της *T* για την οποία ισχύει

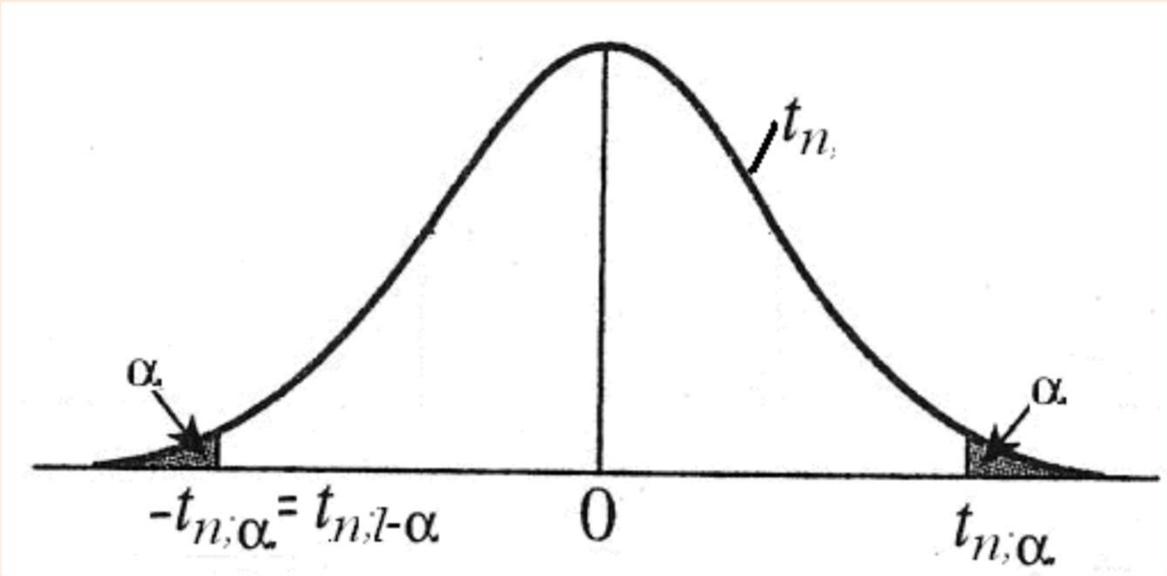
$$P(T > t_{n;\alpha}) = \alpha, \text{ ή ισοδύναμα, } P(T \leq t_{n;\alpha}) = 1-\alpha$$

Η κατανομή t ή κατανομή Student



Η κατανομή t ή κατανομή Student

Όπως στην *κανονική κατανομή*, λόγω συμμετρίας της γραφικής παράστασης της *συνάρτησης πυκνότητας* της t_n , προφανώς ισχύει ότι: $t_{n;1-\alpha} = - t_{n;\alpha}$.

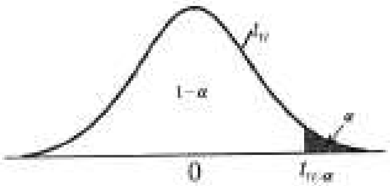


Τιμές $t_{n;a}$ της κατανομής t_n

Ο Πίνακας δίνει τα άνω α -ποσοστιαία σημεία

της κατανομής t με n βαθμούς ελευθερίας

Αν $T \sim t_n$, ***ισχύει,*** $P(T > t_{n;\alpha}) = \alpha$. ***Επίσης, ισχύει,*** $t_{n;1-\alpha} = -t_{n;\alpha}$



n	*α = 0.1*	*α = 0.05*	*α = 0.025*	*α = 0.01*	*α = 0.005*
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
∞	1.282	1.645	1.960	2.326	2.576

για παράδειγμα:

$t_{8;0,10} = 1,397$

επίσης

$t_{8;0,90} = -t_{8;0,10} = -1,397$