



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Μεθοδολογία των Επιστημών του Ανθρώπου: Στατιστική

Ενότητα 2: Επαγωγική Στατιστική

Βασίλης Γιαλαμάς
Σχολή Επιστημών της Αγωγής
Τμήμα Εκπαίδευσης και Αγωγής στην Προσχολική
Ηλικία

Περιεχόμενα ενότητας

Παρουσιάζονται οι βασικές έννοιες του ελέγχου υποθέσεων που αφορούν στη σύγκριση μέσων τιμών και τη συνάφεια μεταξύ δυο μεταβλητών. Γίνεται επίσης εισαγωγή στην εκτιμητική της μέσης τιμής με τη βοήθεια των διαστημάτων εμπιστοσύνης.



Δειγματοληψία (1 από 2)

- Στην Επαγωγική στατιστική οδηγούμαστε σε συμπεράσματα και αποφάσεις για τις παραμέτρους ενός πληθυσμού με τη βοήθεια ενός **τυχαίου δείγματος** που επιλέγεται από τον πληθυσμό με τη διαδικασία της απλής τυχαίας δειγματοληψίας.
- Τυχαίο είναι ένα δείγμα όταν όλα τα άτομα του πληθυσμού έχουν την ίδια πιθανότητα να επιλεγούν σ' αυτό.
- Βασικές μέθοδοι απλής τυχαίας δειγματοληψίας:
 - Κλήρωση
 - Πίνακας τυχαίων αριθμών



Δειγματοληψία (2 από 2)

- Αν ένα **στατιστικό**, δηλ. μια ποσότητα που μπορεί να υπολογιστεί από τα στοιχεία του δείγματος (π.χ. μέση τιμή, τυπική απόκλιση), υπολογιστεί σ' όλα τα δυνατά τυχαία δείγματα του ίδιου μεγέθους N που μπορεί να ληφθούν από ένα πληθυσμό δημιουργείται μια κατανομή τιμών του που ονομάζεται **δειγματοληπτική κατανομή του στατιστικού**.
- **Δειγματοληπτική κατανομή της μέσης τιμής** είναι η κατανομή των μέσων τιμών όλων των τυχαίων δειγμάτων ίδιου μεγέθους N που μπορεί να ληφθούν από ένα πληθυσμό.



Παράδειγμα δειγματοληπτικής κατανομής (1 από 5)

Θα μελετήσουμε την κατανομή της ηλικίας σε υποθετικό πληθυσμό από 5 παιδιά. Οι ηλικίες των παιδιών είναι: $X_1 = 4$, $X_2 = 6$, $X_3 = 8$, $X_4 = 10$, $X_5 = 12$.

Η μέση τιμή του πληθυσμού είναι ίση με:

$$\mu = \frac{\sum_{i=1}^5 X_i}{5} = \frac{4+6+8+10+12}{5} = 8$$

και η διακύμανσή του υπολογίζεται:

$$\sigma^2 = \frac{\sum_{i=1}^5 (X_i - \mu)^2}{5} = \frac{(-4)^2 + (-2)^2 + (2)^2 + (4)^2}{5} = 8$$



Παράδειγμα δειγματοληπτικής κατανομής (2 από 5)

Ας σχηματίσουμε τώρα όλα τα $(5 \times 5 = 25)$ δυνατά δείγματα μεγέθους $N=2$ από τον πληθυσμό των 5 παιδιών. Αυτά φαίνονται στον δίπλα πίνακα:

	4	6	8	10	12
4	4,4 (4)	4,6 (5)	4,8 (6)	4,10 (7)	4,12 (8)
6	6,4 (5)	6,6 (6)	6,8 (7)	6,10 (8)	6,12 (9)
8	8,4 (6)	8,6 (7)	8,8 (8)	8,10 (9)	8,12 (10)
10	10,4(7)	10,6(8)	10,8 (9)	10,10 (10)	10,12 (11)
12	12,4(8)	12,6(9)	12,8 (10)	12,10 (11)	12,12 (12)



Παράδειγμα δειγματοληπτικής κατανομής (3 από 5)

Δίπλα δίνεται ο πίνακας της δειγματοληπτικής κατανομής συχνοτήτων της μέσης τιμής 25 δειγμάτων μεγέθους $N=2$.

\bar{X}_k	f_k	$\bar{X}_k \cdot f_k$	$(\bar{X}_k - \mu_{\bar{X}})^2 \cdot f_k$
4	1	4	16
5	2	10	18
6	3	18	12
7	4	28	4
8	5	40	0
9	4	36	4
10	3	30	12
11	2	22	18
12	1	12	16
Σύνολο	25	200	100



Παράδειγμα δειγματοληπτικής κατανομής (4 από 5)

Ας υπολογίσουμε τη μέση τιμή και διακύμανση της δειγματοληπτικής κατανομής:

$$\mu_{\bar{X}} = \frac{\sum (\bar{X}_k \cdot f_k)}{n^N} = \frac{200}{25} = 8 = \mu \quad \sigma_{\bar{X}}^2 = \frac{\sum [(\bar{X}_k - \mu_{\bar{X}})^2 \cdot f_k]}{25} = \frac{100}{25} = 4$$

Η $\sigma_{\bar{X}}^2$ δεν είναι ίση με τη διασπορά $\sigma^2 = 8$ του πληθυσμού.

Ισχύει όμως: $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N}$

Η τυπική απόκλιση $\sigma_{\bar{X}} = \sqrt{\sigma_{\bar{X}}^2} = \frac{\sigma}{\sqrt{N}}$ ονομάζεται **τυπικό σφάλμα της μέσης τιμής** ή απλά **τυπικό σφάλμα**.

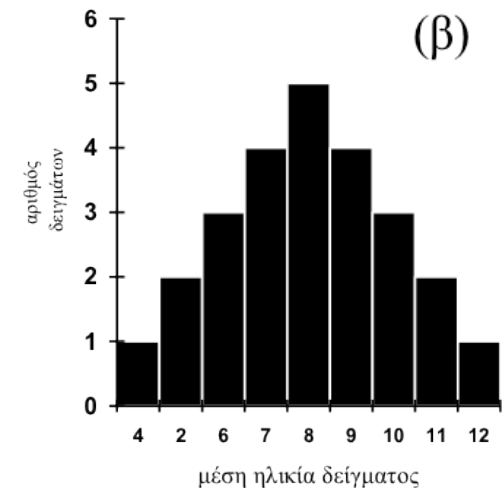
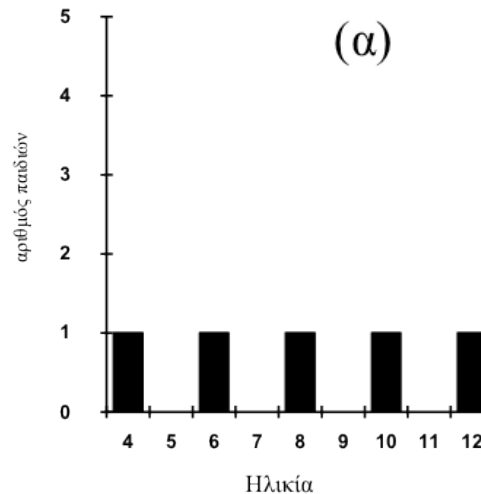
Στο παραπάνω παράδειγμα το τυπικό σφάλμα ισούται με 2.



Παράδειγμα δειγματοληπτικής κατανομής (5 από 5)

Δίπλα δίνεται το ιστόγραμμα των κατανομών του πληθυσμού (α) και το ιστόγραμμα της δειγματοληπτικής κατανομής (β) μέσου όρου δειγμάτων μεγέθους 2.

Είναι εντυπωσιακή η διαφορά ανάμεσα στο ιστόγραμμα της κατανομής του πληθυσμού και το ιστόγραμμα της δειγματοληπτικής κατανομής του. Ενώ στην πρώτη οι τιμές είναι ομοιόμορφα κατανεμημένες, η δεύτερη παρουσιάζει μια κορυφή και είναι απόλυτα συμμετρική.



Κεντρικό οριακό θεώρημα (1 από 2)

Τα αποτελέσματα του προηγούμενου παραδείγματος γενικεύονται με τις δύο παρακάτω προτάσεις:

1. Σε πληθυσμό **κανονικά κατανεμημένο** με μέση τιμή μ , και διακύμανση σ^2 , η δειγματοληπτική κατανομή της μέσης τιμής έχει τις ιδιότητες:

- Είναι κανονική
- $\mu_{\bar{X}} = \mu$
- $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N}$ ($\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$)



Κεντρικό οριακό θεώρημα (2 από 2)

2. Σε πληθυσμό με **οποιοδήποτε τύπο κατανομής**, με μέση τιμή μ και διακύμανση σ^2 , η **δειγματοληπτική κατανομή** της μέσης τιμής έχει τις ιδιότητες:

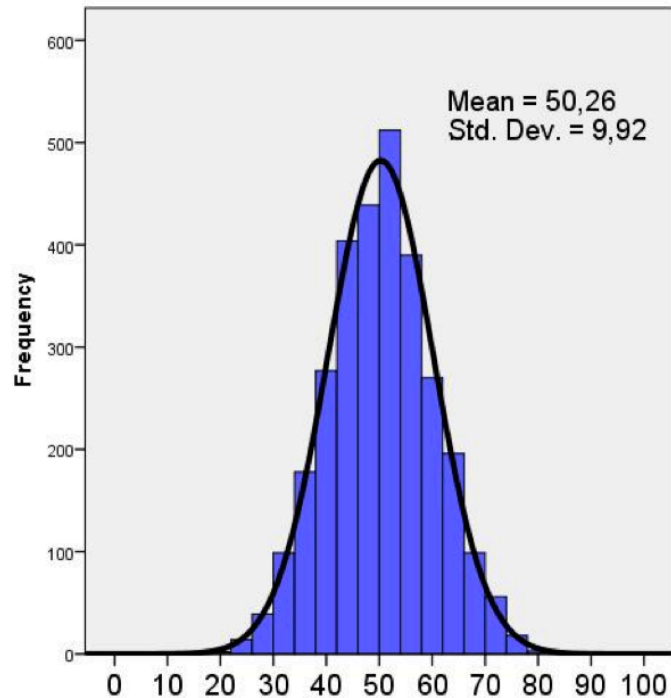
- Είναι κατά προσέγγιση κανονική, όταν το μέγεθος του δείγματος είναι αρκετά μεγάλο ($N \geq 30$).

- $\mu_{\bar{X}} = \mu$

- $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N} \quad (\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}})$



Δειγματοληπτικές κατανομές μέσης τιμής (1 από 8)

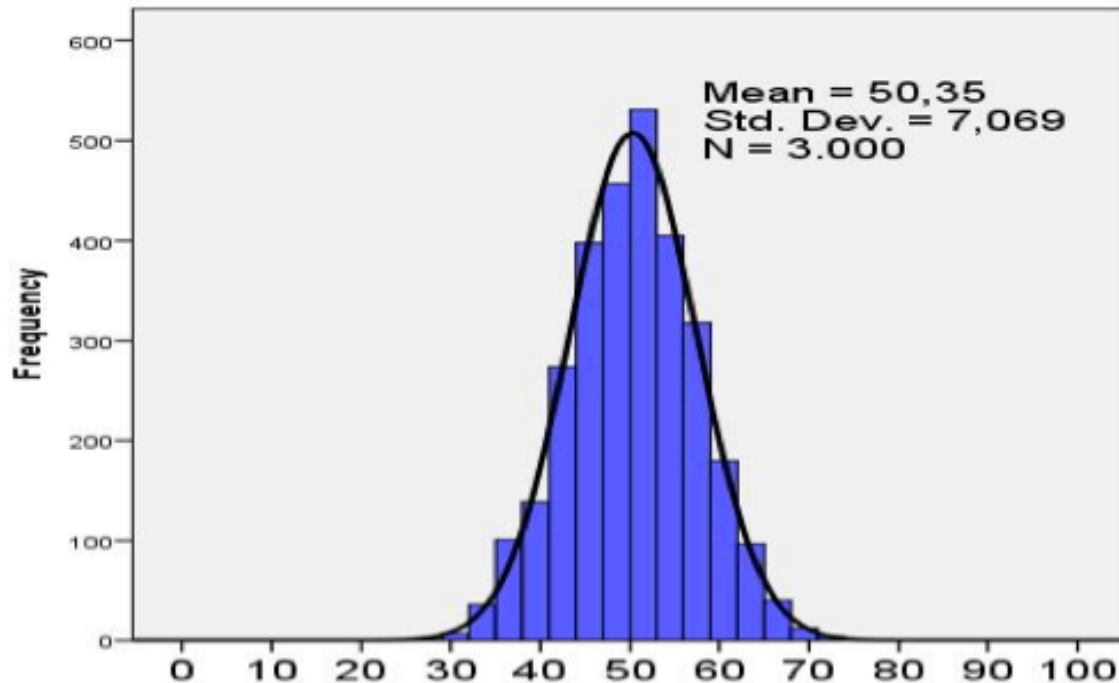


Κατανομή πληθυσμού

Τα δείγματα προέρχονται από πληθυσμό κανονικής κατανομής.



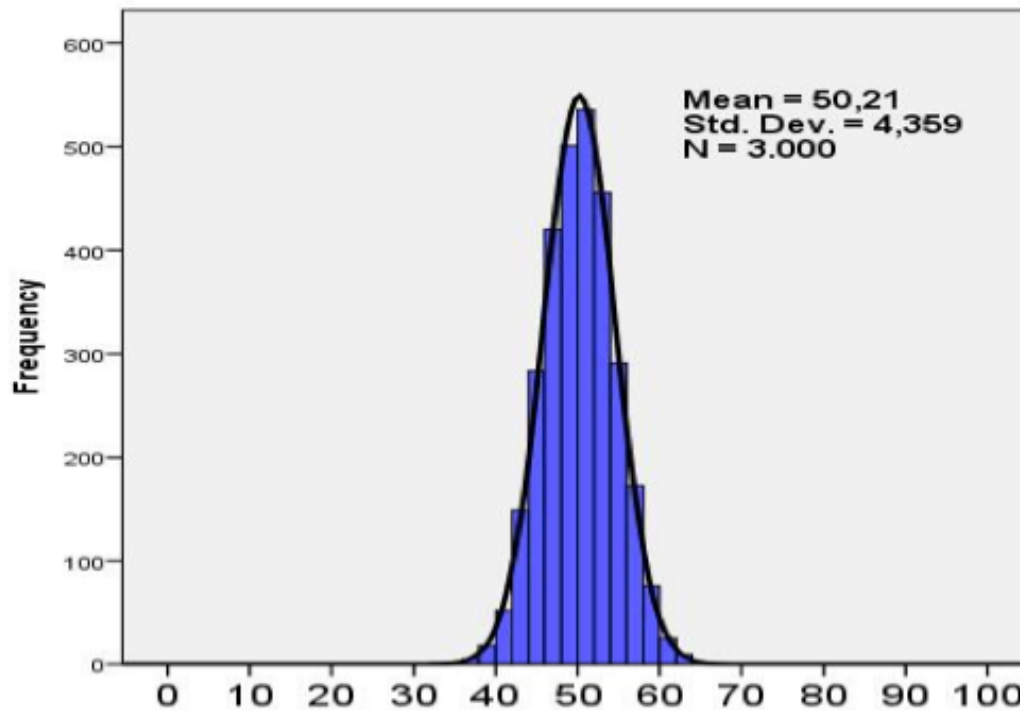
Δειγματοληπτικές κατανομές μέσης τιμής (2 από 8)



Δειγματοληπτική κατανομή μέσης τιμής για $N=2$.



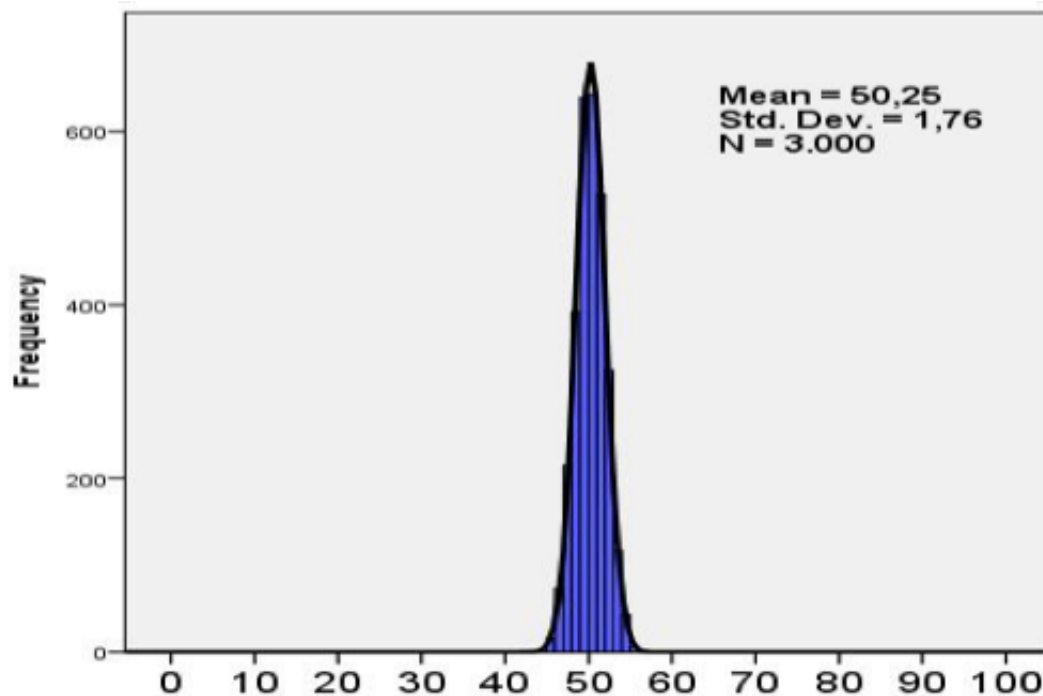
Δειγματοληπτικές κατανομές μέσης τιμής (3 από 8)



Δειγματοληπτική κατανομή μέσης τιμής
για $N=5$.



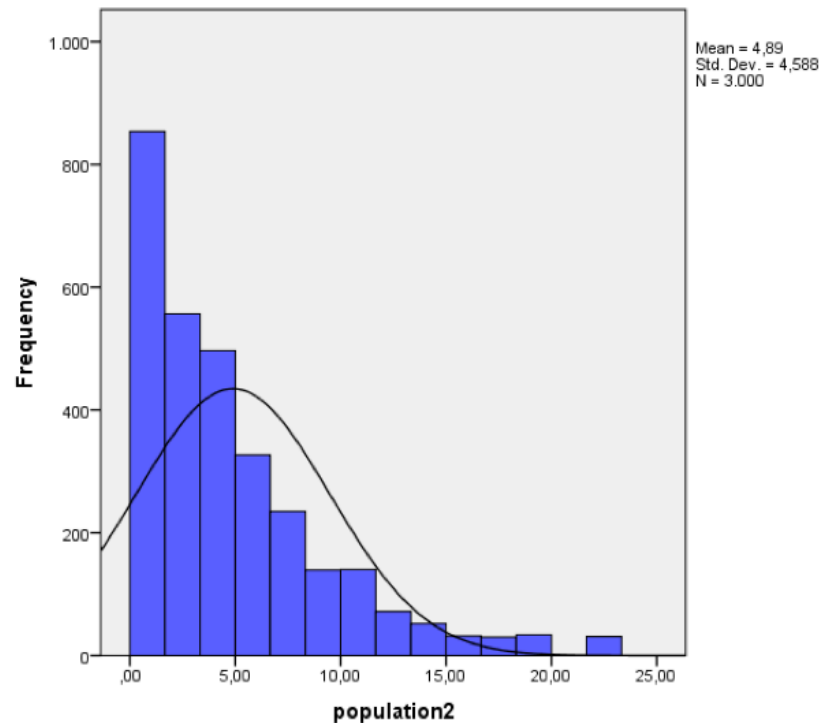
Δειγματοληπτικές κατανομές μέσης τιμής (4 από 8)



Δειγματοληπτική κατανομή μέσης τιμής
για $N=30$.



Δειγματοληπτικές κατανομές μέσης τιμής (5 από 8)

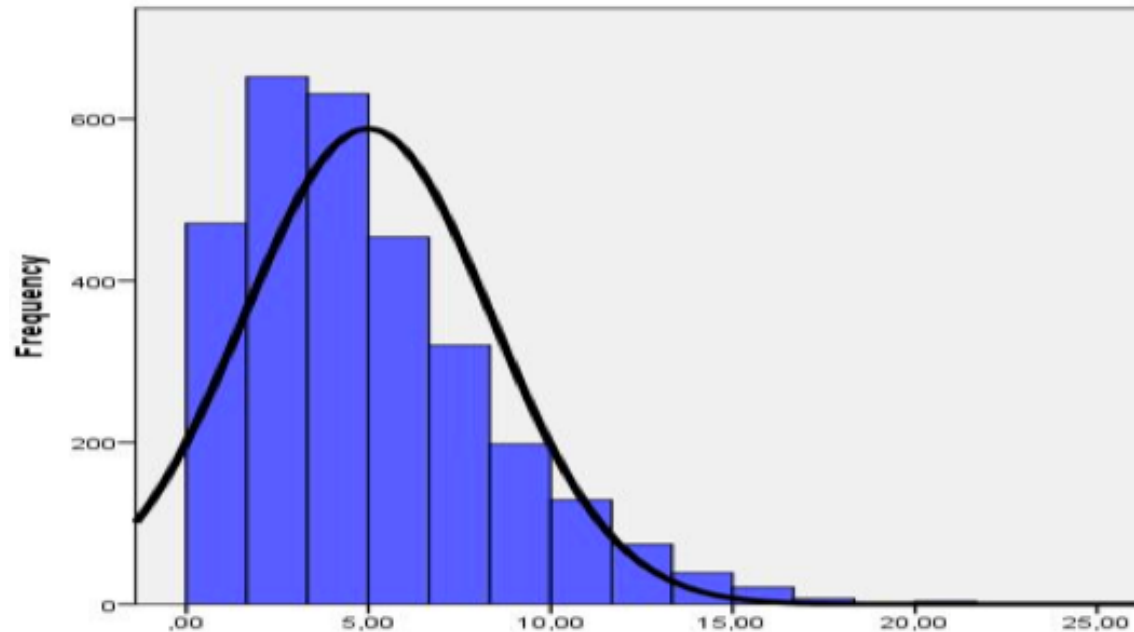


Κατανομή πληθυσμού

Τα δείγματα προέρχονται από πληθυσμό ισχυρής θετικής ασυμμετρίας.



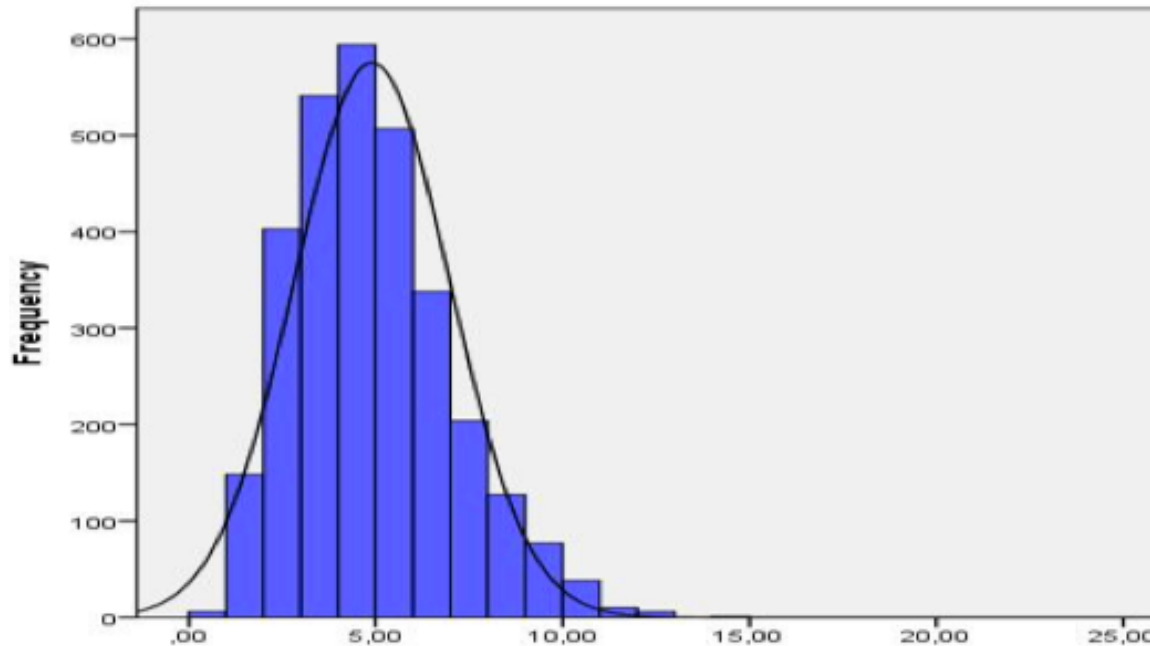
Δειγματοληπτικές κατανομές μέσης τιμής (6 από 8)



Δειγματοληπτική κατανομή μέσης τιμής
για $N=2$.



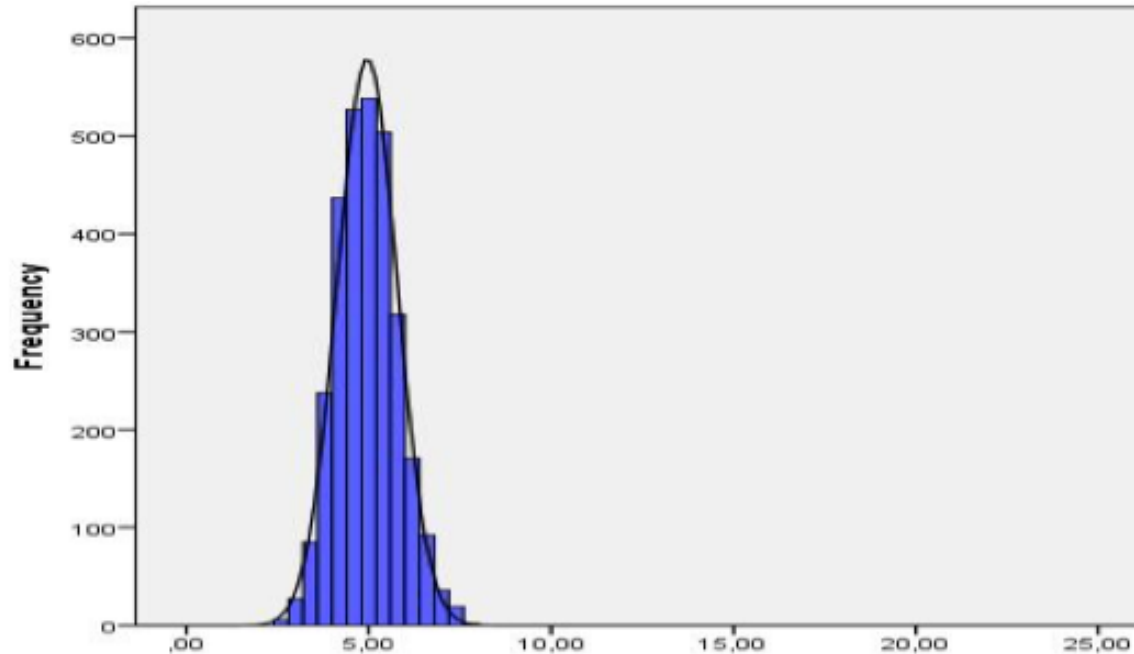
Δειγματοληπτικές κατανομές μέσης τιμής (7 από 8)



Δειγματοληπτική κατανομή μέσης τιμής για $N=5$.



Δειγματοληπτικές κατανομές μέσης τιμής (8 από 8)



Δειγματοληπτική κατανομή μέσης τιμής
για $N=30$.

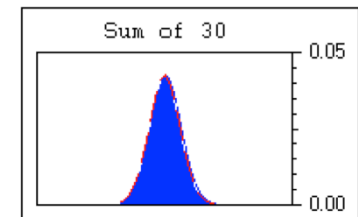
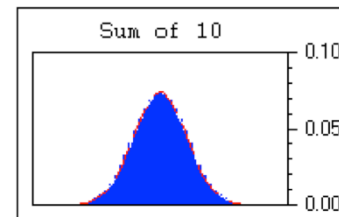
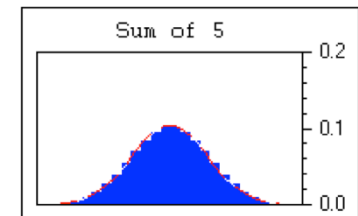
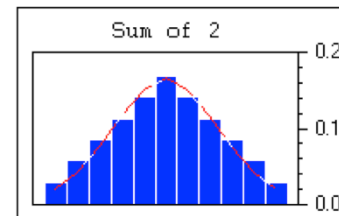
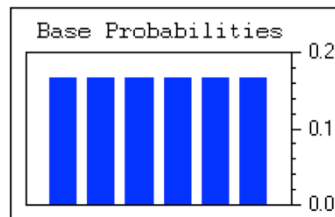


Δειγματοληπτικές κατανομές αθροίσματος (1 από 2)

Δειγματοληπτικές
κατανομές του
αθροίσματος για
 $N=2$, $N=5$, $N=10$ και
 $N=30$.

Τα δείγματα
προέρχονται από
πληθυσμό με
**ομοιόμορφη
κατανομή.**

	probabiliti	
	value	initial su
1	0	0.1667
2	1	0.1667
3	2	0.1667
4	3	0.1667
5	4	0.1667
6	5	0.1667
7		

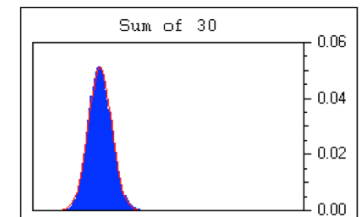
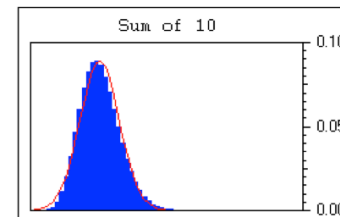
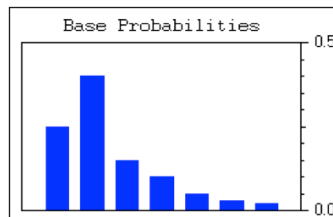
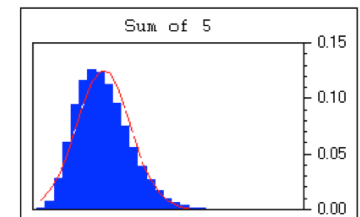
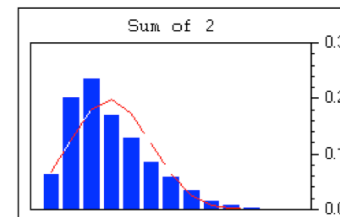


Δειγματοληπτικές κατανομές αθροίσματος (2 από 2)

Δειγματοληπτικές κατανομές του αθροίσματος για $N=2$, $N=5$, $N=10$ και $N=30$.

Τα δείγματα προέρχονται από πληθυσμιακή κατανομή έντονης θετικής λοξότητας.

	probabilities		
	value	initial	sum of 2
1	0	0.2500	0.0625
2	1	0.4000	0.2000
3	2	0.1500	0.2350
4	3	0.1000	0.1700
5	4	0.0500	0.1275
6	5	0.0300	0.0850
7	6	0.0200	0.0590
8			0.0350
9			0.0145
10			0.0070



Δειγματοληπτική κατανομή μέσης τιμής και ΚΟΘ: Άσκηση 3 (1 από 3)

Η επίδοση των μαθητών Γ' γυμνασίου σε ένα τυποποιημένο τεστ μαθηματικών (κλίμακα τιμών 0 έως 20) έχει μέση τιμή $\mu=6,8$ και τυπική απόκλιση $\sigma=3,65$. Με τα δεδομένα αυτά, θα υπολογισθεί η πιθανότητα ενός δείγματος 36 μαθητών Γ' γυμνασίου ενός να επιτύχει μέση τιμή μεγαλύτερη από 10 στο τεστ των Μαθηματικών.



Δειγματοληπτική κατανομή μέσης τιμής και ΚΟΘ: Άσκηση 3 (2 από 3)

Επειδή το δείγμα για το οποίο θέλουμε να εξάγουμε συμπεράσματα είναι «μεγάλο», αφού $N = 36 > 30$, η δειγματοληπτική κατανομή στην οποία ανήκει, σύμφωνα με ΚΟΘ, είναι κανονική.

Για τη μέση τιμή της δειγματοληπτικής κατανομής ισχύει:

$$\mu_{\bar{X}} = \mu = 6,8$$

και η τυπική της απόκλιση είναι: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{3,65}{\sqrt{36}} = 0,61$



Δειγματοληπτική κατανομή μέσης τιμής και ΚΟΘ: Άσκηση 3 (3 από 3)

Για να μπορέσουμε να βρούμε το ποσοστό των τιμών μιας κανονικής κατανομής που είναι μεγαλύτερες από μια δοσμένη τιμή, στην περίπτωση μας 10, πρέπει να μετατρέψουμε τη τιμή αυτή σε τιμή z .

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{10 - 6,8}{0,61} = 5,25$$

Στη συνέχεια ανατρέχοντας στον πίνακα της τυπικής κανονικής κατανομής (Πίνακας Β παραρτήματος) βρίσκουμε τη ζητούμενη πιθανότητα ως εξής:

Αναζητούμε την αναλογία των τιμών που ξεπερνούν την τιμή 5,25. Παρατηρούμε ότι η μεγαλύτερη τιμή στον πίνακα είναι η 5, στην οποία αντιστοιχεί η πιθανότητα 0,0000003. Συμπεραίνουμε, λοιπόν, ότι η πιθανότητα ενός δείγματος 36 μαθητών να έχει μέσο βαθμό πάνω από 10 στο τεστ μαθηματικών, είναι ασήμαντη, δηλαδή μικρότερη από 0,0000003 ή 3 στις 10.000.000 ή 0,00003%.



Τέλος Ενότητας

Επαγωγική Στατιστική

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.



Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστημίων Αθηνών 2015, Βασίλης Γιαλαμάς 2015. Βασίλης Γιαλαμάς. «Μεθοδολογία των Επιστημών του Ανθρώπου: Στατιστική. Επαγωγική Στατιστική». Έκδοση: 1.0. Αθήνα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://opencourses.uoa.gr/courses/ECD102/>.



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.



Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων (1/4)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες

Εικόνα 1, Σελίδα 9: Εικόνα με ιστόγραμμα των κατανομών του πληθυσμού και το ιστόγραμμα της δειγματοληπτικής κατανομής / Copyright Σ. Πατάκης ΑΕΕΔΕ (Εκδόσεις Πατάκη) και Β. Γιαλαμάς, 2004 / Πηγή: «Στατιστικές Τεχνικές και Εφαρμογές στις Επιστήμες της Αγωγής» Β. Γιαλαμάς, Εκδόσεις Πατάκη

Εικόνα 2, Σελίδα 12: Εικόνα ιστογράμματος κατανομής πληθυσμού/ Copyrighted

Εικόνα 3, Σελίδα 13: Εικόνα ιστογράμματος δειγματοληπτικής κατανομής μέσης τιμής με $N=2$ / Copyrighted

Εικόνα 4, Σελίδα 14: Εικόνα ιστογράμματος δειγματοληπτικής κατανομής μέσης τιμής με $N=5$ / Copyrighted



Σημείωμα Χρήσης Έργων Τρίτων (2/4)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες

Εικόνα 5, Σελίδα 15: Εικόνα ιστογράμματος δειγματοληπτικής κατανομής μέσης τιμής με $N=30$ / Copyrighted

Εικόνα 6, Σελίδα 16: Εικόνα ιστογράμματος κατανομής πληθυσμού/ Copyrighted

Εικόνα 7, Σελίδα 17: Εικόνα ιστογράμματος δειγματοληπτικής κατανομής μέσης τιμής με $N=2$ / Copyrighted

Εικόνα 8, Σελίδα 18: Εικόνα ιστογράμματος δειγματοληπτικής κατανομής μέσης τιμής με $N=5$ / Copyrighted



Σημείωμα Χρήσης Έργων Τρίτων (3/4)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες

Εικόνα 9, Σελίδα 19: Εικόνα ιστογράμματος δειγματοληπτικής κατανομής μέσης τιμής με $N=30$ / Copyrighted

Εικόνα 10, Σελίδα 20: Εικόνα ιστογραμμάτων με δειγματοληπτικές κατανομές του αθροίσματος για $N=2$, $N=5$, $N=10$ και $N=30$ (ομοιόμορφη κατανομή) / Copyrighted

Εικόνα 11, Σελίδα 21: Εικόνα ιστογραμμάτων με δειγματοληπτικές κατανομές του αθροίσματος για $N=2$, $N=5$, $N=10$ και $N=30$ (κατανομή έντονης θετικής λοξότητας) / Copyrighted



Σημείωμα Χρήσης Έργων Τρίτων (4/4)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Πίνακες

Πίνακας 1, Σελίδα 6: Πίνακας δειγμάτων / Copyrighted

Πίνακας 2, Σελίδα 7: Πίνακας δειγματοληπτικής κατανομής συχνοτήτων / Copyrighted

