Ασκήσεις Στατιστικής (Διαστήματα Εμπιστοσύνης)

1	**Example 42.1** The average distance measured from a reference point in four independent trials is 2250 m. The mean error of the measuring instrument is $E = 40$ m. Given the confidence level, 95 per cent, find the confidence interval for the quantity measured. Το Ε που σας δίνεται είναι η τυπική απόκλιση και υποθέτουμε κανονική κατανομή.
2	**Example 42.2** The standard deviation of an altimeter is $\sigma = 15$ m. How many altimeters should there be on an airplane so that with confidence level 0.99, the mean error in altitude $\tilde{x}$ is not greater than -30 m., if the errors given by the altimeters are normally distributed, and there are no systematic errors? *altimeter*=όργανο μέτρησης ύψους "*no systematic errors*" εδώ σημαίνει ότι ο μέσος του σφάλματος είναι 0. Υποτίθεται κανονική κατανομή και το σ σας δίνεται. Προσοχή όμως γιατί το διάστημα εμπιστοσύνης που ζητείται είναι μόνο προς την μία κατεύθυνση. Δε θέλουμε η μέτρηση του ύψους (η μέση τιμή των Ν οργάνων) να είναι $\geq Y + 30$ όπου Υ το πραγματικό ύψος. Δε μας πειράζει η μέτρηση του ύψους δηλ. να είναι οσοδήποτε χαμηλή.
3	**15.2.1** Let Φ be the d.f. of the $N(0, 1)$ distribution and let a and b with $a < b$ be such that $\Phi(b) - \Phi(a) = \gamma (0 < \gamma < 1)$. Show that $b - a$ is minimum if $b = c\ (> 0)$ and $a = -c$. (See also the discussion of the second part of Example 1.) Μια απλή ιδιότητα της κανονικής κατανομής. Το μόνο που χρειάζεστε για να το δείξετε είναι ότι η κανονική πυκνότητα είναι άρτια συνάρτηση, είναι αύξουσα μέχρι το 0 και φθίνουσα μετά το 0.
4	**15.2.8** Consider the independent random samples $X_1, \ldots, X_m$ from $N(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_n$ from $N(\mu_2, \sigma_2^2)$, where σ_1, σ_2 are known and μ_1, μ_2 are unknown, and let the r.v. $T_{m,n}(\mu_1 - \mu_2)$ be defined by $T_{m,n}(\mu_1 - \mu_2) = \frac{(\overline{X}_m - \overline{Y}_n) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/m) + (\sigma_2^2/n)}}.$ Then show that: **i)** A confidence interval for $\mu_1 - \mu_2$, based on $T_{m,n}(\mu_1 - \mu_2)$, with confidence coefficient $1 - \alpha$ is given by $\left[(\overline{X}_m - \overline{Y}_n) - b\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}},\ (\overline{X}_m - \overline{Y}_n) - a\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}\right],$ where a and b are such that $\Phi(b) - \Phi(a) = 1 - \alpha$; **ii)** The shortest confidence interval of the aforementioned form is provided by the last expression above with $-a = b = z_{\frac{\alpha}{2}}$.

5	**9.1.6** If the r.v. X is distributed as $U(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, show that the r.v. $Y = \tan X$ is distributed as Cauchy. Also find the distribution of the r.v. $Z = \sin X$.
6	**Example 3.1** A random sample of size 6 is drawn from a $N(\mu, 12)$ distribution. Find $P(2.76 < S^2 < 22.2)$. Χρησιμοποιείστε την κατανομή χ^2.
7	If $W_1, \ldots, W_k$ are independent $\chi^2_{\nu_1}, \ldots, \chi^2_{\nu_k}$ then $\sum_{i=1}^{k} W_i \sim \chi^2_{\nu}$ Αποδείξτε αυτό.
8	**Theorem 4.6.1** Suppose $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, \ldots, n$ and that they are independent random variables. Let $Y = (\sum_i a_i X_i) + b$ for some constants $\{a_i\}$ and b. Then $$Y \sim N\left((\sum_i a_i \mu_i) + b, \sum_i a_i^2 \sigma_i^2\right).$$
9	**Theorem 4.6.2** Suppose $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, \ldots, n$ and also that the $\{X_i\}$ are *independent*. Let $U = \sum_{i=1}^{n} a_i X_i$ and $V = \sum_{i=1}^{n} b_i X_i$ for some constants $\{a_i\}$ and $\{b_i\}$. Then $\text{Cov}(U, V) = \sum_i a_i b_i \sigma_i^2$. Εδώ Cov(U,Y) είναι η συνδιακύμανση (covariance) των U, V. Αν δε θυμάστε πώς ορίζεται ψάξτε το.
10	Define bias in terms of expected value. Is it possible for a statistic to be unbiased yet very imprecise? How about being very accurate but biased? Why is a 99% confidence interval wider than a 95% confidence interval? When you construct a 95% confidence interval, what are you 95% confident about? What is the effect of sample size on the width of a confidence interval?
11	You take a sample of 22 from a population of test scores, and the mean of your sample is 60. (a) You know the standard deviation of the population is 10. What is the 99% confidence interval on the population mean. (b) Now assume that you do not know the population standard deviation, but the standard deviation in your sample is 10. What is the 99% confidence interval on the mean now?
12	You read about a survey in a newspaper and find that 70% of the 250 people sampled prefer Candidate A. You are surprised by this survey because you thought that more like 50% of the population preferred this candidate. Based on this sample, is 50% a possible population proportion? Compute the 95% confidence interval to be sure.