

Συγκριτική στατιστική μεθοδολογία
(γνωστή και ως)
Στατιστική συμπερασματολογία
(Inferential statistics)

Βήματα που ακολουθούμε

Βήμα	Γενική μεθοδολογία	Συγκεκριμένο παράδειγμα
1	Θέτουμε το επίπεδο σημαντικότητας	The level of significance was set at 0.05 (two-sided)
2	Θέτουμε την ερώτηση και με βάση τη φύση των δεδομένων (εξαρτημένης και ανεξάρτητης μεταβλητής) αποφασίζουμε τι στατιστικό τεστ θα χρησιμοποιήσουμε.	Θέλουμε να ξέρουμε σε ένα δείγμα 40 ανδρών (m) και 60 γυναικών (w) αν ο ΔΜΣ διαφέρει στατιστικά σημαντικά μεταξύ των δύο ομάδων. Εφόσον ο ΔΜΣ ακολουθεί την κανονική κατανομή τότε θα κάνουμε t-test.
3	Θέτουμε τη μηδενική υπόθεση (H_0). Συνήθως σαν H_0 θέτουμε αυτό που θα θέλαμε να απορρίψουμε.	$H_0: \mu_m = \mu_w$ ή $\mu_m - \mu_w = 0$. Οι μέσες τιμές του ΔΜΣ ανδρών και των γυναικών δεν διαφέρουν μεταξύ τους.
4	Θέτουμε την εναλλακτική υπόθεση (H_1). Είναι αυτή που θα αποδεχτούμε σε κάποια όρια εμπιστοσύνης έχοντας απορρίψει τη μηδενική υπόθεση,	$H_1: \mu_m \neq \mu_w$ ή $\mu_m - \mu_w \neq 0$. Επειδή δεν μπορούμε να ξέρουμε εκ των προτέρων ποιος ΔΜΣ είναι μεγαλύτερος γι' αυτό και θα πραγματοποιήσουμε το δίπλευρο τεστ ($\mu_m > \mu_w$ και $\mu_m < \mu_w$).
5	Διενεργούμε το τεστ και με βάση την τιμή p που μας δίνει τότε: Α) Αν $p < 0.05$ τότε απορρίπτουμε τη μηδενική υπόθεση Β) Δηλώνουμε ότι δεν έχουμε λόγους να απορρίψουμε τη μηδενική υπόθεση.	Α) 28.7 ± 5.1 vs 26.6 ± 4.9 , $t(98) = 2.1$, $p = 0.042$ ($p < 0.05$) Η μέση τιμή ΔΜΣ των ανδρών είναι στατιστικά σημαντικά μεγαλύτερη από αυτή των γυναικών (με πιθανότητα να έχουμε κάνει λάθος $p = 0.05$ – λάθος τύπου I) Β) 28.7 ± 5.3 vs 26.9 ± 5.1 , $t(98) = 1.7$, $p = 0.092$ ($p = NS$) Ο ΔΜΣ δεν διαφέρει μεταξύ των δύο φύλων (εμπεριέχει την άγνωστη πιθανότητα να έχουμε κάνει λάθος τύπου II).

Εξαρτημένη μεταβλητή

Συνεχής που ακολουθεί την κανονική κατανομή

Όταν υπάρχει σχέση αλλά όχι εξάρτηση

Συντελεστής συσχέτισης του Pearson
(Pearson's correlation coefficient - r)

Correlations			
		APACHE II score	SOFA score
APACHE II score	Pearson Correlation	1	.529**
	Sig. (2-tailed)		.000
	N	1251	1250
SOFA score	Pearson Correlation	.529**	1
	Sig. (2-tailed)	.000	
	N	1250	1284

** . Correlation is significant at the 0.01 level (2-tailed).

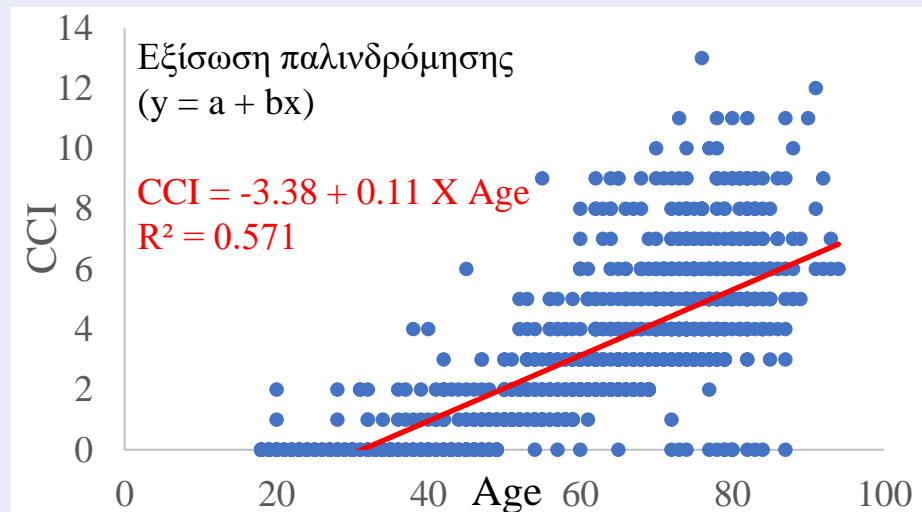
The correlation between APACHE II score and SOFA score is statistically significant ($r=0.529$, $p<0.01$).
Σημείωση: Όταν οι συνεχείς μεταβλητές δεν ακολουθούν την κανονική κατανομή τότε αναφέρουμε το Spearman's rho correlation coefficient.

Ανεξάρτητη μεταβλητή

Συνεχής που ακολουθεί την κανονική κατανομή

Όταν υπάρχει εξάρτηση (συνάρτηση)

Γραμμική παλινδρόμηση
(Linear regression)



Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-3.380	.181		-18.662	.000
	Age (years)	.108	.003	.756	37.839	.000

a. Dependent Variable: Charlons's comorbidity index

Περίπτωση όταν έχουμε δύο μετρήσεις της ίδιας μεταβλητής που είναι συνεχής και ακολουθεί την κανονική κατανομή. Κλασική περίπτωση πειραματικών σχεδιασμών ΠΡΙΝ-ΜΕΤΑ.

t-test για ζευγαρωτές μετρήσεις
Paired samples t-test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PCT day 1	11.2428	102	36.01335	3.56585
	PCT day 7	2.6761	102	9.53918	.94452
Pair 2	suPAR day 1	18.6320	43	8.73439	1.33198
	suPAR day 7	20.0631	43	17.93557	2.73515

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	PCT day 1 & PCT day 7	102	.376	.000
Pair 2	suPAR day 1 & suPAR day 7	43	.780	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PCT day 1 - PCT day 7	8.56676	33.60883	3.32777	1.96536	15.16817	2.574	101	.011
Pair 2	suPAR day 1 - suPAR day 7	-1.43106	12.39360	1.89001	-5.24525	2.38313	-.757	42	.453

There was a significant decrease of the procalcitonin levels from day 1 to day 7 ($t(101)=2.57$, $p=0.011$). With regards to suPAR levels there was no analogous decrease.

Περίπτωση όταν έχουμε δύο μετρήσεις της ίδιας μεταβλητής που είναι συνεχής και **ΔΕΝ** ακολουθεί την κανονική κατανομή. Κλασική περίπτωση πειραματικών σχεδιασμών ΠΡΙΝ-ΜΕΤΑ.

Wilcoxon test για ζευγαρωτές μετρήσεις
Wilcoxon Signed Ranks Test (Wilcoxon Paired samples Signed Ranks Test)

		Ranks		
		N	Mean Rank	Sum of Ranks
PCT day 7 - PCT day 1	Negative Ranks	79	53.97	4264.00
	Positive Ranks	20	34.30	686.00
	Ties	3		
	Total	102		
suPAR day 7 - suPAR day 1	Negative Ranks	24	20.88	501.00
	Positive Ranks	19	23.42	445.00
	Ties	0		
	Total	43		

Test Statistics ^a		
	PCT day 7 - PCT day 1	suPAR day 7 - suPAR day 1
Z	-6.244 ^b	-.338 ^b
Asymp. Sig. (2-tailed)	.000	.735

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

There was a significant decrease of the procalcitonin levels from day 1 to day 7 (Wilcoxon test, $p < 0.01$). With regards to suPAR levels there was no analogous decrease.

Παραμετρικά τεστ: Πραγματοποιούνται όταν οι συνεχείς μεταβλητές ακολουθούν την κανονική κατανομή
Μη παραμετρικά τεστ: Πραγματοποιούνται όταν οι συνεχείς μεταβλητές δεν ακολουθούν την κανονική κατανομή και σε αρκετές περιπτώσεις με ταξινομικές μεταβλητές.

Τι ψάχνουμε	Παραμετρικό τεστ	Μη παραμετρικό τεστ
Συσχέτιση	Pearson's correlation coefficient	Spearman's correlation coefficient
Μεταβολή	Paired samples t-test	Wilcoxon test

Εξαρτημένη μεταβλητή

Συνεχής που ακολουθεί την κανονική κατανομή

Ανεξάρτητη μεταβλητή

Κατηγορική διωνυμική μεταβλητή

t-test για ανεξάρτητα δείγματα (t-test for independent samples)

Group Statistics

	Patient's gender	N	Mean	Std. Deviation	Std. Error Mean
APACHE II score	Male	806	20.1687	8.04684	.28344
	Female	421	20.1188	8.19147	.39923

Table 1. Baseline Characteristics of Enrolled Patients

Characteristic	Placebo (N = 78)	BCG (N = 72)	p value
Age, years, mean (SD)	79.6 (7.8)	79.9 (7.6)	0.802
Male gender, no. (%)	35 (44.9)	32 (44.4)	1.000
Charlson's Comorbidity Index, mean (SD)	5.5 (1.9)	5.5 (2.2)	0.909
APACHE II score on study enrolment, mean (SD)	7.9 (3.0)	8.1 (2.9)	0.701
SOFA score on study enrolment, mean (SD)	1.2 (1.4)	1.0 (1.1)	0.586

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
APACHE II score	Equal variances assumed	.008	.930	.103	1225	.918	.04997	.48688	-.90524	1.00518
	Equal variances not assumed			.102	838.908	.919	.04997	.48961	-.91104	1.01098

Εξαρτημένη μεταβλητή	Ανεξάρτητη μεταβλητή
Συνεχής που δεν ακολουθεί την κανονική κατανομή (ή ταξινομική μεταβλητή)	Κατηγορική διωνυμική μεταβλητή
Mann-Whitney Test (μη παραμετρικό τεστ)	

Ranks					Test Statistics ^a	
	Patient's gender	N	Mean Rank	Sum of Ranks	suPAR day 1	
suPAR day 1	Male	234	168.89	39519.50	Mann-Whitney U	12024.500
	Female	121	195.62	23670.50	Wilcoxon W	39519.500
	Total	355			Z	-2.327
					Asymp. Sig. (2-tailed)	.020

a. Grouping Variable: Patient's gender

Table 1. Baseline Characteristics of Patients.

	Standard-of-care (N=130)	Standard-of-care + Anakinra (N=130)	P-Value
Age, years, median (Q ₁ -Q ₃)	63 (53-73)	64 (52-74)	0.982
Male sex, no. (%)	84 (64.6)	81 (62.3)	0.797
Severity indexes, median (Q ₁ -Q ₃)			
Charlson's Comorbidity Index	2.5 (1.0-4.0)	3.0 (1.0-4.0)	0.864
APACHE II score	6.0 (4.0-9.0)	6.5 (4.8-9.0)	0.962
SOFA score	2.0 (1.0-3.0)	2.0 (1.0-3.0)	0.975
Pneumonia Severity Index	68.5 (54.8-82.0)	67.5 (51.8-83.0)	0.507

Παραμετρικά τεστ: Πραγματοποιούνται όταν οι συνεχείς μεταβλητές ακολουθούν την κανονική κατανομή
Μη παραμετρικά τεστ: Πραγματοποιούνται όταν οι συνεχείς μεταβλητές δεν ακολουθούν την κανονική κατανομή και σε αρκετές περιπτώσεις με ταξινομικές μεταβλητές.

Τι ψάχνουμε	Παραμετρικό τεστ	Μη παραμετρικό τεστ
Συσχέτιση	Pearson's correlation coefficient	Spearman's correlation coefficient
Μεταβολή (Πριν – Μετά)	Paired samples t-test	Wilcoxon test
Διαφορά μεταξύ δύο ομάδων	Independent samples t-test	Mann-Whitney test
Διαφορά μεταξύ πολλών ομάδων (>2)*	ANOVA (Ανάλυση διασποράς)	Kruskal-Wallis tests
Επαναλαμβανόμενες μετρήσεις (>2)*	ANOVA for repeated measures	Friedman test

* - Σε περίπτωση που υπάρχει στατιστική σημαντικότητα τότε πραγματοποιούμε post-hoc pairwise tests (με ή χωρίς corrections, συνήθως κατά Bonferroni), για να αναδείξουμε ποια ζευγάρια ομάδων ή μετρήσεων διαφέρουν στατιστικά σημαντικά μεταξύ τους.

Εξαρτημένη μεταβλητή

Ανεξάρτητη μεταβλητή

Κατηγορική διωνυμική μεταβλητή

Συνεχής ανεξαρτήτως κατανομής, ή ταξινομική ή κατηγορική μεταβλητή (με προϋποθέσεις).

Λογιστική παλινδρόμηση (Binary logistic regression)

Dependent Variable Encoding

Original Value	Internal Value
Reconstitution	0
Death	1

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	APACHE II score	.073	.008	82.973	1	.000	1.076	1.059	1.093
	Constant	-1.990	.179	124.160	1	.000	.137		

a. Variable(s) entered on step 1: APACHE II score.

OR=1.08, 95% CI=1.06-1.09, p<0.01

Στην περίπτωση που γνωρίζουμε και το χρόνο του συμβάντος (εδώ του θανάτου)
Μοντέλα επιβίωσης (Cox regression)

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
								Lower	Upper
	APACHE II score	.064	.006	125.754	1	.000	1.066	1.054	1.078

HR=1.07, 95% CI=1.05-1.08, p<0.01

Ειδική περίπτωση

Εξαρτημένη μεταβλητή

Ανεξάρτητη μεταβλητή

Κατηγορική διωνυμική μεταβλητή

Κατηγορική διωνυμική μεταβλητή

Λογιστική παλινδρόμηση (Binary logistic regression) +
Πίνακες συνάφειας 2Χ2 (2Χ2 contingency tables) with Fisher exact test

Presence of acute respiratory distress syndrome * 28-day outcome
Crosstabulation

Chi-Square Tests

		28-day outcome		Total	
		Reconstitution	Death		
Presence of acute respiratory distress syndrome	No	Count	462	205	667
		% within Presence of acute respiratory distress syndrome	69.3%	30.7%	100.0%
	Yes	Count	329	292	621
		% within Presence of acute respiratory distress syndrome	53.0%	47.0%	100.0%
Total	Count	791	497	1288	
	% within Presence of acute respiratory distress syndrome	61.4%	38.6%	100.0%	

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	35.995 ^a	1	.000		
Continuity Correction ^b	35.311	1	.000		
Likelihood Ratio	36.134	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	35.967	1	.000		
N of Valid Cases	1288				

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Presence of acute respiratory distress syndrome (No / Yes)	2.000	1.593	2.512
For cohort 28-day outcome = Reconstitution	1.307	1.195	1.430
For cohort 28-day outcome = Death	.654	.568	.753
N of Valid Cases	1288		

Fisher exact test, $p < 0.01$, OR=2.0 95% CI= 1.6-2.5

Ειδική περίπτωση

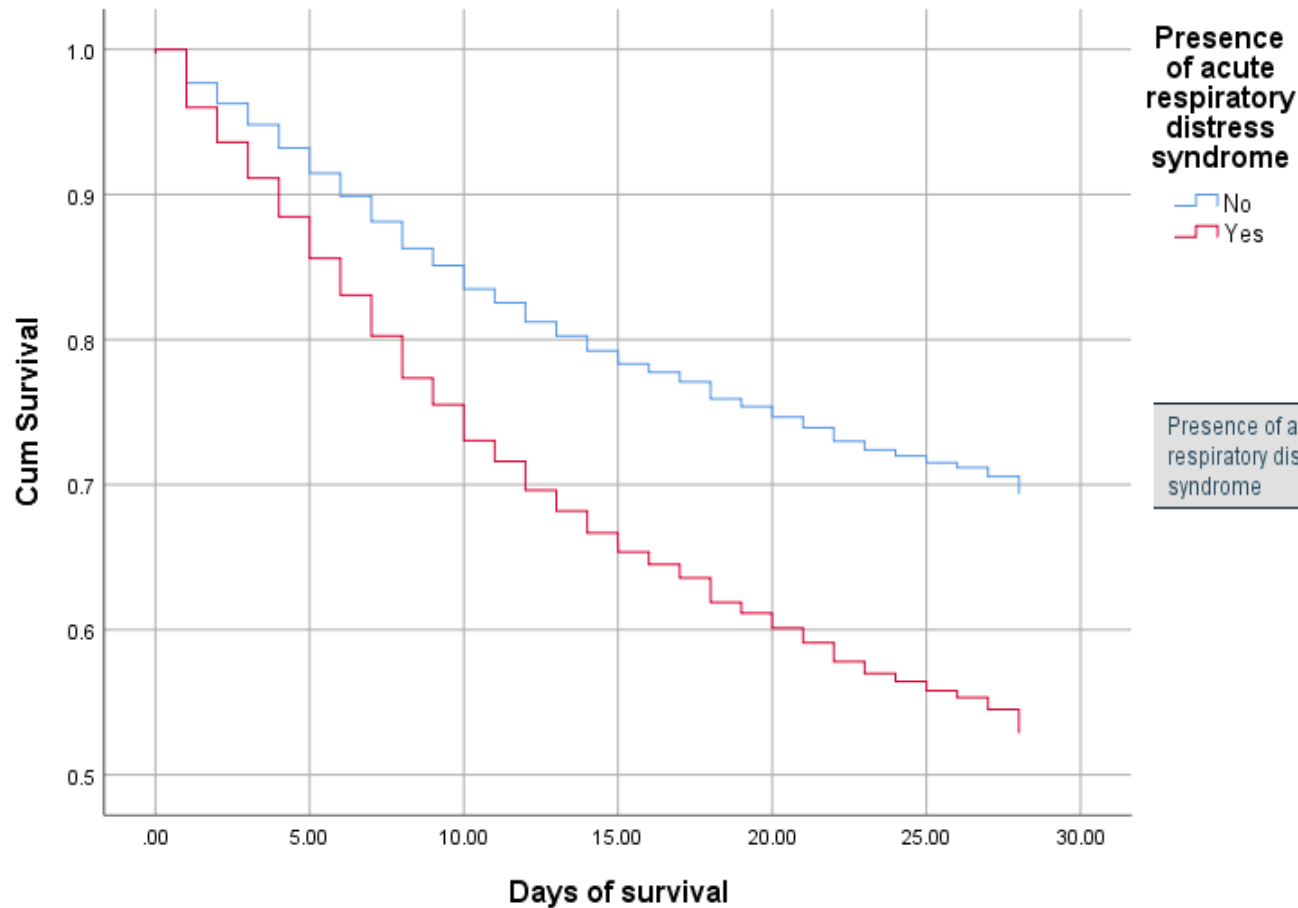
Εξαρτημένη μεταβλητή

Ανεξάρτητη μεταβλητή

Κατηγορική διωνυμική μεταβλητή

Κατηγορική διωνυμική μεταβλητή

Όταν είναι γνωστός ο χρόνος του συμβάντος
Μοντέλα επιβίωσης (Cox regression) με γράφημα επιβίωσης



Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Presence of acute respiratory distress syndrome	.554	.092	36.392	1	.000	1.740	1.454	2.084

HR=1.7 95% CI=1.5-2.1, $p < 0.01$