9.15. CONFRONTO TRA DUE DISTRIBUZIONI OSSERVATE: IL METODO DI KOLMOGOROV-SMIRNOV PER 2 CAMPIONI INDIPENDENTI CON DATI ORDINALI DISCRETI O A GRUPPI E CON DATI CONTINUI

Una tabella $\mathbf{2 \times N}$, in cui siano riportate le classi di frequenza di due distribuzioni osservate, può essere analizzata mediante il test di Kolmogorov-Smirnov. A differenza del test χ^2 e del test G, come condizione di validità non richiede che il numero minimo di osservazioni sia **N > 30** e che in ogni classe sia $n_i \geq 5$.

Il test di Kolmogorov-Smirnov, necessario per piccoli campioni, è utile anche per grandi campioni, quando si abbia un numero elevato di classi od intervalli, poiché in tali condizioni ha una potenza maggiore del test chi quadrato e del test G. Ovviamente **la potenza massima è con dati in una scala continua**, che rappresenta la **proposta originale** di **Kolmogorov** e **Smirnov**.

Le distribuzioni di frequenza possono riguardare qualunque variabile, da quelle classiche di peso ed altezza, per le quali possono essere applicati anche test parametrici, Alle misure espresse come **rapporti**, **percentuali**, **valori angolari**, **colorazioni**, **indici**, **punteggi assoluti o relativi**, ecc., purché esse possano essere tradotte in ranghi, cioè ordinate per dimensioni o intensità.

Le indicazioni bibliografiche sono uguali a quelle già riportate nel test di Kolmogorov-Smirnov per un campione. Differiscono le indicazioni per le tabelle dei valori critici, che ovviamente per il metodo ora illustrato devono considerare una casistica più complessa, in particolare se i due campioni non sono bilanciati.

Il test di Kolmogorov-Smirnov per due campioni indipendenti è utilizzato per verificare l'ipotesi alternativa se le distribuzioni di frequenza di due campioni appartengano a popolazioni differenti.

E' un test generalista, cioè permette di valutare la significatività complessiva dovuta a differenze

- sia nella tendenza centrale,
- sia nella dispersione,
- sia nella simmetria e,
- sebbene in modo meno evidente, nella curtosi.

Non è un test specifico per nessuno di questi fattori. Quindi non deve essere utilizzato se l'ipotesi verte su un parametro specifico, come la media e la varianza oppure la simmetria. Tuttavia non tutti i parametri pesano nello stesso modo; è più sensibile alle differenze nelle tendenze centrali, perché incidono sulla differenza complessiva tra le due distribuzioni in modo più marcato.

Se il test risulta significativo, per individuare con esattezza quale caratteristica della distribuzione determini la differenza riscontrata, occorre di conseguenza ricorrere anche all'uso di altri test, che ne verifichino solamente una e che, ovviamente, per quell'uso specifico sono più potenti.

Sotto l'aspetto della ricerca applicata, è utile quando si intende verificare se due serie di valori possono o meno appartenere alla stessa popolazione. Infatti, se hanno origine diversa, è logico supporre che le due serie di dati campionari differiscano in almeno un parametro, senza che a priori sia noto quale.

Il test di Kolmogorov-Smirnov per due campioni può essere utilizzato

- sia con dati misurati su una **scala ordinale discreta** o con dati continui **raggruppati in classi**,
- sia con **dati continui** di una **scala di rapporti oppure a intervalli oppure ordinale**.

PER DATI DISCRETI O RAGGRUPPATI

Tra i testi internazionali, questo metodo è riportato in

- Siegel **Sidney** e N. John jr. **Castellan** del 1988 (***Nonparametric Statistics for the Behavioral Sciences***, (McGraw-Hill, London), tradotto in italiano nel 1992 ***Statistica non parametrica*** 2° ed., McGraw-Hill Libri Italia, Milano, 472 pp.)

del quale sono seguite le indicazioni nella presentazione del metodo e della logica.

In modo analogo al test per un campione, questo per 2 campioni richiede

- che dapprima sia effettuata la trasformazione delle frequenze assolute in frequenze relative entro ogni campione, mediante il rapporto della frequenza di ogni classe con il numero totale di osservazioni;
- successivamente, che entro gli stessi intervalli sia attuato il confronto tra le due frequenze cumulate, per quantificare la deviazione o differenza massima.

Sulla base dell'**ipotesi** formulata, se unilaterale oppure bilaterale, la differenza massima può essere considerata con il segno oppure in valore assoluto.

Indicando con $O_1(X_i)$ ogni valore della sommatoria dei dati osservati nel primo campione e con $O_2(X_i)$ ogni valore della sommatoria dei dati osservati nel secondo campione,

- nel caso di un **test ad una coda** si deve calcolare la deviazione massima D con il segno

$$D = \text{diff. mass. } (O_1(Xi) - O_2(Xi))$$

- per un **test a due code** non è importante conoscere la direzione della differenza; lo scarto massimo è quindi calcolato in valore assoluto

$$D = \text{diff. mass. } | O_1(Xi) - O_2(Xi) |$$

Nel caso di **piccoli campioni**, quando le due distribuzioni hanno al massimo 25 osservazioni (altri testi definiscono i campioni come piccoli fino ad un massimo di 40 osservazioni), si può ricorrere a tabelle

specifiche per verificare se la differenza massima tra le cumulate delle frequenze relative supera il valore critico e quindi sia significativa. Sulle tabelle di significatività, le proposte in letteratura sono numerose, in quanto questo test è stato tra quelli che hanno suscitato un dibattito scientifico maggiore. Quelle riportate in queste dispense sono tra le più semplici.

Il **valore da confrontare (J)** è ottenuto **moltiplicando la differenza massima D per le dimensioni dei due campioni n_1 e n_2**.

$$\mathbf{J} = D \cdot n_1 \cdot n_2$$

I **valori critici** sono differenti

- per **test a una coda**, riportati nella prima tabella
- per **test a due code**, riportati nella seconda tabella.

La loro impostazione è uguale. Per il loro uso, ricordare che

- sulla prima riga si trova il numero di osservazioni del primo (n_1) campione e sulla prima colonna il numero di osservazioni del secondo (n_2) campione,
- alla loro intersezione si trovano i tre valori **J** critici in colonna, associati rispettivamente dall'alto al basso alla probabilità α **= 0.10**, alla probabilità α **= 0.05** e a quella α **= 0.01**.

Il valore **J** calcolato (**J = D· $n_1 \cdot n_2$**) indica una differenza significativa quando è uguale o superiore a quello critico riportato nella tabella.

Nel caso di un **test ad una coda**, per esempio

- con 10 osservazioni nel campione 1 e con 12 osservazioni nel campione 2, la differenza tra le due distribuzioni cumulate è significativa
- alla probabilità $\alpha = 0.10$ quando $J \geq 52$,
- alla probabilità $\alpha = 0.05$ quando $J \geq 60$,
- alla probabilità $\alpha = 0.01$ quando $J \geq 74$.

E' possibile osservare che la distribuzione dei valori critici è simmetrica. Alle stesse probabilità sono identici, quando si hanno 12 osservazioni nel campione 1 e 10 osservazioni nel campione 2.

Valori critici (J) nel test, <u>ad una coda</u>, di Kolmogorov-Smirnov per 2 campioni indipendenti.
Il valore superiore è per $\alpha = 0.10$; quello centrale per $\alpha = 0.05$ e quello inferiore per $\alpha = 0.01$.

	n_2																						
n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
3	9	10	11	15	15	16	21	19	22	24	25	26	30	30	32	36	36	37	42	40	43	45	46
	9	**10**	**13**	**15**	**16**	**19**	**21**	**22**	**25**	**27**	**28**	**31**	**33**	**34**	**35**	**39**	**40**	**41**	**45**	**46**	**47**	**51**	**52**
	**	**	**	**	**19**	**22**	**27**	**28**	**31**	**33**	**34**	**37**	**42**	**43**	**43**	**48**	**49**	**52**	**54**	**55**	**58**	**63**	**64**
4	10	16	13	16	18	24	21	24	26	32	29	32	34	40	37	40	41	48	45	48	49	56	53
	10	**16**	**16**	**18**	**21**	**24**	**25**	**28**	**29**	**36**	**33**	**38**	**38**	**44**	**44**	**46**	**49**	**52**	**52**	**56**	**57**	**60**	**61**
	**	**	**17**	**22**	**25**	**32**	**29**	**34**	**37**	**40**	**41**	**46**	**46**	**52**	**53**	**56**	**57**	**64**	**64**	**66**	**69**	**76**	**73**
5	11	13	20	19	21	23	26	30	30	32	35	37	45	41	44	46	47	55	51	54	56	58	65
	13	**16**	**20**	**21**	**24**	**26**	**28**	**35**	**35**	**36**	**40**	**42**	**50**	**46**	**49**	**51**	**56**	**60**	**60**	**62**	**65**	**67**	**75**
	**	**17**	**25**	**26**	**29**	**33**	**36**	**40**	**41**	**46**	**48**	**51**	**60**	**56**	**61**	**63**	**67**	**75**	**75**	**76**	**81**	**82**	**90**
6	15	16	19	24	24	26	30	32	33	42	37	42	45	48	49	54	54	56	60	62	63	72	67
	15	**18**	**21**	**30**	**25**	**30**	**33**	**36**	**38**	**48**	**43**	**48**	**51**	**54**	**56**	**66**	**61**	**66**	**69**	**70**	**73**	**78**	**78**
	**	**22**	**26**	**36**	**31**	**38**	**42**	**44**	**49**	**54**	**54**	**60**	**63**	**66**	**68**	**78**	**77**	**80**	**84**	**88**	**91**	**96**	**96**
7	15	18	21	24	35	28	32	34	38	40	44	49	48	51	54	56	59	61	70	68	70	72	74
	16	**21**	**24**	**25**	**35**	**34**	**36**	**40**	**43**	**45**	**50**	**56**	**56**	**58**	**61**	**64**	**68**	**72**	**77**	**77**	**79**	**83**	**85**
	19	**25**	**29**	**31**	**42**	**42**	**46**	**50**	**53**	**57**	**59**	**70**	**70**	**71**	**75**	**81**	**85**	**87**	**98**	**97**	**99**	**103**	**106**
8	16	24	23	26	28	40	33	40	41	48	47	50	52	64	57	62	64	72	71	74	76	88	81
	19	**24**	**26**	**30**	**34**	**40**	**40**	**44**	**48**	**52**	**53**	**58**	**60**	**72**	**65**	**72**	**73**	**80**	**81**	**84**	**89**	**96**	**95**
	22	**32**	**33**	**38**	**42**	**48**	**49**	**56**	**59**	**64**	**66**	**72**	**75**	**88**	**81**	**88**	**91**	**100**	**100**	**106**	**107**	**120**	**118**
9	21	21	26	30	32	33	45	43	45	51	51	54	60	61	65	72	70	73	78	79	82	87	88
	21	**25**	**28**	**33**	**36**	**40**	**54**	**46**	**51**	**57**	**57**	**63**	**69**	**68**	**74**	**81**	**80**	**83**	**90**	**91**	**94**	**99**	**101**
	27	**29**	**36**	**42**	**46**	**49**	**63**	**61**	**62**	**69**	**73**	**77**	**84**	**86**	**92**	**99**	**99**	**103**	**111**	**111**	**117**	**123**	**124**
10	19	24	30	32	34	40	43	50	48	52	55	60	65	66	69	72	74	90	80	86	88	92	100
	22	**28**	**35**	**36**	**40**	**44**	**46**	**60**	**57**	**60**	**62**	**68**	**75**	**76**	**77**	**82**	**85**	**100**	**91**	**98**	**101**	**106**	**110**
	28	**34**	**40**	**44**	**50**	**56**	**61**	**70**	**69**	**74**	**78**	**84**	**90**	**94**	**97**	**104**	**104**	**120**	**118**	**120**	**125**	**130**	**140**
11	22	26	30	33	38	41	45	48	66	54	59	63	66	69	72	76	79	84	85	99	95	98	100
	25	**29**	**35**	**38**	**43**	**48**	**51**	**57**	**66**	**64**	**67**	**72**	**76**	**80**	**83**	**87**	**92**	**95**	**101**	**110**	**108**	**111**	**116**
	31	**37**	**41**	**49**	**53**	**59**	**62**	**69**	**88**	**77**	**85**	**89**	**95**	**100**	**104**	**108**	**114**	**117**	**124**	**143**	**132**	**138**	**143**
12	24	32	32	42	40	48	51	52	54	72	61	68	72	76	77	84	85	92	93	98	100	108	106
	27	**36**	**36**	**48**	**45**	**52**	**57**	**60**	**64**	**72**	**71**	**78**	**84**	**88**	**89**	**96**	**98**	**104**	**108**	**110**	**113**	**132**	**120**
	33	**40**	**46**	**54**	**57**	**64**	**69**	**74**	**77**	**96**	**92**	**94**	**102**	**108**	**111**	**120**	**121**	**128**	**132**	**138**	**138**	**156**	**153**
13	25	29	35	37	44	47	51	55	59	61	78	72	75	79	81	87	89	95	97	100	105	109	111
	28	**33**	**40**	**43**	**50**	**53**	**57**	**62**	**67**	**71**	**91**	**78**	**86**	**90**	**94**	**98**	**102**	**108**	**112**	**117**	**120**	**124**	**131**
	34	**41**	**48**	**54**	**59**	**66**	**73**	**78**	**85**	**92**	**104**	**102**	**106**	**112**	**118**	**121**	**127**	**135**	**138**	**143**	**150**	**154**	**160**
14	26	32	37	42	49	50	54	60	63	68	72	84	80	84	87	92	94	100	112	108	110	116	119
	31	**38**	**42**	**48**	**56**	**58**	**63**	**68**	**72**	**78**	**78**	**98**	**92**	**96**	**99**	**104**	**108**	**114**	**126**	**124**	**127**	**132**	**136**
	37	**46**	**51**	**60**	**70**	**72**	**77**	**84**	**89**	**94**	**102**	**112**	**111**	**120**	**124**	**130**	**135**	**142**	**154**	**152**	**157**	**164**	**169**
15	30	34	45	45	48	52	60	65	66	72	75	80	90	87	91	99	100	110	111	111	111	123	130
	33	**38**	**50**	**51**	**56**	**60**	**69**	**75**	**76**	**84**	**86**	**92**	**105**	**101**	**105**	**111**	**113**	**125**	**126**	**130**	**134**	**141**	**145**
	42	**46**	**60**	**63**	**70**	**75**	**84**	**90**	**95**	**102**	**106**	**111**	**135**	**120**	**130**	**138**	**142**	**150**	**156**	**160**	**165**	**174**	**180**
16	30	40	41	48	51	64	61	66	69	76	79	84	87	112	94	100	104	112	114	118	122	136	130
	34	**44**	**46**	**54**	**58**	**72**	**68**	**76**	**80**	**88**	**90**	**96**	**101**	**112**	**109**	**116**	**120**	**128**	**130**	**136**	**140**	**152**	**148**
	43	**52**	**56**	**66**	**71**	**88**	**86**	**94**	**100**	**108**	**112**	**120**	**120**	**144**	**139**	**142**	**149**	**156**	**162**	**168**	**174**	**184**	**185**
17	32	37	44	49	54	57	65	69	72	77	81	87	91	94	119	102	108	113	118	122	128	132	137
	35	**44**	**49**	**56**	**61**	**65**	**74**	**77**	**83**	**89**	**94**	**99**	**105**	**109**	**136**	**118**	**125**	**130**	**135**	**141**	**146**	**150**	**156**
	43	**53**	**61**	**68**	**75**	**81**	**92**	**97**	**104**	**111**	**118**	**124**	**130**	**139**	**153**	**150**	**157**	**162**	**168**	**175**	**181**	**187**	**192**
18	36	40	46	54	56	62	72	72	76	84	87	92	99	100	102	126	116	120	126	128	133	144	142
	39	**46**	**51**	**66**	**64**	**72**	**51**	**82**	**87**	**96**	**98**	**104**	**111**	**116**	**118**	**144**	**127**	**136**	**144**	**148**	**151**	**162**	**161**
	48	**56**	**63**	**78**	**81**	**88**	**99**	**104**	**108**	**120**	**121**	**130**	**138**	**142**	**150**	**180**	**160**	**170**	**177**	**184**	**189**	**198**	**201**
19	36	41	47	54	59	64	70	74	79	85	89	94	100	104	108	116	133	125	128	132	137	142	148
	40	**49**	**56**	**61**	**68**	**73**	**80**	**85**	**92**	**98**	**102**	**108**	**113**	**120**	**125**	**127**	**152**	**144**	**147**	**151**	**159**	**162**	**168**
	49	**57**	**67**	**77**	**85**	**91**	**99**	**104**	**114**	**121**	**127**	**135**	**142**	**149**	**157**	**160**	**190**	**171**	**183**	**189**	**197**	**204**	**211**
20	37	48	55	56	61	72	73	90	84	92	95	100	110	112	113	120	125	140	134	138	143	152	155
	41	**52**	**60**	**66**	**72**	**80**	**83**	**100**	**95**	**104**	**108**	**114**	**125**	**128**	**130**	**136**	**144**	**160**	**154**	**160**	**163**	**172**	**180**
	52	**64**	**75**	**80**	**87**	**100**	**103**	**120**	**117**	**128**	**135**	**142**	**150**	**156**	**162**	**170**	**171**	**200**	**193**	**196**	**203**	**212**	**220**
21	42	45	51	60	70	71	78	80	85	93	97	112	111	114	118	126	128	134	147	142	147	156	158
	45	**52**	**60**	**69**	**77**	**81**	**90**	**91**	**101**	**108**	**112**	**126**	**126**	**130**	**135**	**144**	**147**	**154**	**168**	**163**	**170**	**177**	**182**
	54	**64**	**75**	**84**	**98**	**100**	**111**	**118**	**124**	**132**	**138**	**154**	**156**	**162**	**168**	**177**	**183**	**193**	**210**	**205**	**212**	**222**	**225**
22	40	48	54	62	68	74	79	86	99	98	100	108	111	118	122	128	132	138	142	176	151	158	163
	46	**56**	**62**	**70**	**77**	**84**	**91**	**98**	**110**	**110**	**117**	**124**	**130**	**136**	**141**	**148**	**151**	**160**	**163**	**198**	**173**	**182**	**188**
	55	**66**	**76**	**88**	**97**	**106**	**111**	**120**	**143**	**138**	**143**	**152**	**160**	**168**	**175**	**184**	**189**	**196**	**205**	**242**	**217**	**228**	**234**
23	43	49	56	63	70	76	82	88	95	100	105	110	117	122	128	133	137	143	147	151	184	160	169
	47	**57**	**65**	**73**	**79**	**89**	**94**	**101**	**108**	**113**	**120**	**127**	**134**	**140**	**146**	**151**	**159**	**163**	**170**	**173**	**207**	**183**	**194**
	58	**69**	**81**	**91**	**99**	**107**	**117**	**125**	**132**	**138**	**150**	**157**	**165**	**174**	**181**	**189**	**197**	**203**	**212**	**217**	**253**	**228**	**242**
24	45	56	58	72	72	88	87	92	98	108	109	116	123	136	132	144	142	152	156	158	160	192	178
	51	**60**	**67**	**78**	**83**	**96**	**99**	**106**	**111**	**132**	**124**	**132**	**141**	**152**	**150**	**162**	**162**	**172**	**177**	**182**	**183**	**216**	**204**
	63	**76**	**82**	**96**	**103**	**120**	**123**	**130**	**138**	**156**	**154**	**164**	**174**	**184**	**187**	**198**	**204**	**212**	**222**	**228**	**228**	**264**	**254**
25	46	53	65	67	74	81	88	100	100	106	111	119	130	130	137	142	148	155	158	163	169	178	200
	52	**61**	**75**	**78**	**85**	**95**	**101**	**110**	**116**	**120**	**131**	**136**	**145**	**148**	**156**	**161**	**168**	**180**	**182**	**188**	**194**	**204**	**225**
	64	**73**	**90**	**96**	**106**	**118**	**124**	**140**	**143**	**153**	**160**	**169**	**180**	**185**	**192**	**201**	**211**	**220**	**225**	**234**	**242**	**254**	**275**

Valori critici nel test, a due code, di Kolmogorov-Smirnov per 2 campioni indipendenti.
Il valore superiore è per $\alpha = 0.10$; quello centrale per $\alpha = 0.05$ e quello inferiore per $\alpha = 0.01$.

n_1	n_2 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1																			19	20	21	22	23	24	25
2					10	12	14	16	18	18	20	22	24	24	26	28	30	32	32	34	36	38	38	40	42
								16	**18**	**20**	**22**	**24**	**26**	**26**	**28**	**30**	**32**	**34**	**36**	**38**	**38**	**40**	**42**	**44**	**46**
																			38	**40**	**42**	**44**	**46**	**48**	**50**
3			9	12	15	15	18	21	21	24	27	27	30	33	33	36	36	39	42	42	45	48	48	51	54
					15	**18**	**21**	**21**	**24**	**27**	**30**	**30**	**33**	**36**	**36**	**39**	**42**	**45**	**45**	**48**	**51**	**51**	**54**	**57**	**60**
									27	**30**	**33**	**36**	**39**	**42**	**42**	**45**	**48**	**51**	**54**	**57**	**57**	**60**	**63**	**66**	**69**
4			12	16	16	18	21	24	27	28	29	36	35	38	40	44	44	46	49	52	52	56	57	60	63
			0	**16**	**20**	**20**	**24**	**29**	**28**	**30**	**33**	**36**	**39**	**42**	**44**	**48**	**48**	**50**	**53**	**60**	**59**	**62**	**64**	**68**	**68**
						24	**28**	**32**	**36**	**36**	**40**	**44**	**48**	**48**	**52**	**56**	**60**	**60**	**64**	**68**	**72**	**72**	**76**	**80**	**84**
5		10	15	16	20	24	25	27	30	35	35	36	46	42	50	48	50	52	56	60	60	63	65	67	75
			15	**20**	**25**	**24**	**28**	**30**	**35**	**40**	**39**	**43**	**45**	**46**	**55**	**54**	**55**	**60**	**61**	**65**	**69**	**70**	**72**	**76**	**80**
					25	**30**	**35**	**35**	**40**	**45**	**45**	**50**	**52**	**56**	**60**	**64**	**68**	**70**	**71**	**80**	**80**	**83**	**87**	**90**	**95**
6		12	15	18	24	30	28	30	33	36	38	48	46	48	51	54	56	66	64	66	69	70	73	78	78
			18	**20**	**24**	**30**	**30**	**34**	**39**	**40**	**43**	**48**	**52**	**54**	**57**	**60**	**62**	**72**	**70**	**72**	**75**	**78**	**80**	**90**	**88**
				24	**30**	**36**	**36**	**40**	**45**	**48**	**54**	**60**	**60**	**64**	**69**	**72**	**73**	**84**	**83**	**88**	**90**	**92**	**97**	**102**	**107**
7		14	18	21	25	28	35	34	36	40	44	46	50	56	56	59	61	65	69	72	77	77	80	84	86
			21	**24**	**28**	**30**	**42**	**40**	**42**	**46**	**48**	**53**	**56**	**63**	**62**	**64**	**68**	**72**	**76**	**79**	**91**	**84**	**89**	**92**	**97**
				28	**35**	**36**	**42**	**48**	**49**	**53**	**59**	**60**	**65**	**77**	**75**	**77**	**84**	**87**	**91**	**93**	**105**	**103**	**108**	**112**	**115**
8		16	21	24	27	30	34	40	40	44	48	52	54	58	60	72	68	72	74	50	81	84	89	96	95
		16	**21**	**28**	**30**	**34**	**40**	**48**	**46**	**48**	**53**	**60**	**62**	**64**	**67**	**80**	**77**	**80**	**82**	**88**	**89**	**94**	**98**	**104**	**104**
				32	**35**	**40**	**48**	**56**	**55**	**60**	**64**	**68**	**72**	**76**	**81**	**88**	**88**	**94**	**98**	**104**	**107**	**112**	**115**	**128**	**125**
9		18	21	27	30	33	36	40	54	50	52	57	59	63	69	69	74	81	80	84	90	91	94	99	101
		18	**24**	**28**	**35**	**39**	**42**	**46**	**54**	**53**	**59**	**63**	**65**	**70**	**75**	**78**	**82**	**90**	**89**	**93**	**99**	**101**	**106**	**111**	**114**
			27	**36**	**40**	**45**	**49**	**55**	**63**	**63**	**70**	**75**	**78**	**84**	**90**	**94**	**99**	**108**	**107**	**111**	**117**	**122**	**126**	**132**	**135**
10		18	24	28	35	36	40	44	50	60	57	60	64	68	75	76	79	82	85	100	95	98	101	106	110
		20	**27**	**30**	**40**	**40**	**46**	**48**	**53**	**70**	**60**	**66**	**70**	**74**	**80**	**84**	**89**	**92**	**94**	**110**	**105**	**108**	**114**	**118**	**125**
			30	**36**	**45**	**48**	**53**	**60**	**63**	**80**	**77**	**80**	**84**	**90**	**100**	**100**	**106**	**108**	**113**	**130**	**126**	**130**	**137**	**140**	**150**
11		20	27	29	35	38	44	48	52	57	66	64	67	73	76	80	85	88	92	96	101	110	108	111	117
		22	**30**	**33**	**39**	**43**	**48**	**53**	**59**	**60**	**77**	**72**	**75**	**82**	**84**	**89**	**93**	**97**	**102**	**107**	**112**	**121**	**119**	**124**	**129**
			33	**40**	**45**	**54**	**59**	**64**	**70**	**77**	**88**	**86**	**91**	**96**	**102**	**106**	**110**	**118**	**122**	**127**	**134**	**143**	**142**	**150**	**154**
12		22	27	36	36	48	46	52	57	60	64	72	71	78	84	88	90	96	99	104	108	110	113	132	120
		24	**30**	**36**	**43**	**48**	**53**	**60**	**63**	**66**	**72**	**84**	**81**	**86**	**93**	**96**	**100**	**108**	**108**	**116**	**120**	**124**	**125**	**144**	**138**
			36	**44**	**50**	**60**	**60**	**68**	**75**	**80**	**86**	**96**	**95**	**104**	**108**	**116**	**119**	**126**	**130**	**140**	**141**	**148**	**149**	**168**	**165**
13		24	30	35	40	46	50	54	59	64	67	71	91	78	87	91	96	99	104	108	113	117	120	125	131
		26	**33**	**39**	**45**	**52**	**56**	**62**	**65**	**70**	**75**	**81**	**91**	**89**	**96**	**101**	**105**	**110**	**114**	**120**	**126**	**130**	**135**	**140**	**145**
			39	**48**	**52**	**60**	**65**	**72**	**78**	**84**	**91**	**95**	**117**	**104**	**115**	**121**	**127**	**131**	**138**	**143**	**150**	**156**	**161**	**166**	**172**
14		24	33	38	42	48	56	58	63	68	73	78	78	98	92	96	100	104	110	114	126	124	127	132	136
		26	**36**	**42**	**46**	**54**	**63**	**64**	**70**	**74**	**82**	**86**	**89**	**112**	**98**	**106**	**111**	**116**	**121**	**126**	**140**	**138**	**142**	**146**	**150**
			42	**48**	**56**	**64**	**77**	**76**	**84**	**90**	**96**	**104**	**104**	**126**	**123**	**126**	**134**	**140**	**148**	**152**	**161**	**164**	**170**	**176**	**182**
15		26	33	40	50	51	56	60	69	75	76	84	87	92	105	101	105	111	114	125	126	130	134	141	145
		28	**36**	**44**	**55**	**57**	**62**	**67**	**75**	**80**	**84**	**93**	**96**	**98**	**120**	**114**	**116**	**123**	**127**	**135**	**138**	**144**	**149**	**156**	**160**
			42	**52**	**60**	**69**	**75**	**81**	**90**	**100**	**102**	**108**	**115**	**123**	**135**	**133**	**142**	**147**	**152**	**160**	**168**	**173**	**179**	**186**	**195**
16		28	36	44	48	54	59	72	69	76	80	88	91	96	101	112	109	116	120	128	130	136	141	152	149
		30	**39**	**48**	**54**	**60**	**64**	**80**	**78**	**84**	**89**	**96**	**101**	**106**	**114**	**128**	**124**	**128**	**133**	**140**	**145**	**150**	**157**	**168**	**167**
			45	**56**	**64**	**72**	**77**	**88**	**94**	**100**	**106**	**116**	**121**	**126**	**133**	**160**	**143**	**154**	**160**	**168**	**173**	**180**	**187**	**200**	**199**
17		30	36	44	50	56	61	68	74	79	85	90	96	100	105	109	136	118	126	132	136	142	146	151	156
		32	**42**	**48**	**55**	**62**	**68**	**77**	**82**	**89**	**93**	**100**	**105**	**111**	**116**	**124**	**136**	**133**	**141**	**146**	**151**	**157**	**163**	**168**	**173**
			48	**60**	**68**	**73**	**84**	**88**	**99**	**106**	**110**	**119**	**127**	**134**	**142**	**143**	**170**	**164**	**166**	**175**	**180**	**187**	**196**	**203**	**207**
18		32	39	46	52	66	65	72	81	82	88	96	99	104	111	116	118	144	133	136	144	148	152	162	162
		34	**45**	**50**	**60**	**72**	**72**	**80**	**90**	**92**	**97**	**108**	**110**	**116**	**123**	**128**	**133**	**162**	**142**	**152**	**159**	**164**	**170**	**180**	**180**
			51	**60**	**70**	**84**	**87**	**94**	**108**	**108**	**118**	**126**	**131**	**140**	**147**	**154**	**164**	**180**	**176**	**182**	**189**	**196**	**204**	**215**	**216**
19	19	32	42	49	56	64	69	74	80	85	92	99	104	110	114	120	126	133	152	144	147	152	159	164	168
		36	**45**	**53**	**61**	**70**	**76**	**82**	**89**	**94**	**102**	**108**	**114**	**121**	**127**	**133**	**141**	**142**	**171**	**160**	**163**	**169**	**177**	**183**	**187**
		38	**54**	**64**	**71**	**83**	**91**	**98**	**107**	**113**	**122**	**130**	**138**	**148**	**152**	**160**	**166**	**176**	**190**	**187**	**199**	**204**	**209**	**218**	**224**
20	20	34	42	52	60	66	72	80	84	100	96	104	108	114	125	128	132	136	144	160	154	160	164	172	180
		38	**48**	**60**	**65**	**72**	**79**	**88**	**93**	**110**	**107**	**116**	**120**	**126**	**135**	**140**	**146**	**152**	**160**	**180**	**173**	**176**	**184**	**192**	**200**
		40	**57**	**68**	**80**	**88**	**93**	**104**	**111**	**130**	**127**	**140**	**143**	**152**	**160**	**168**	**175**	**182**	**187**	**220**	**199**	**212**	**219**	**228**	**235**
21	21	36	45	52	60	69	77	81	90	95	101	108	113	126	126	130	136	144	147	154	168	163	171	177	182
		38	**51**	**59**	**69**	**75**	**91**	**89**	**99**	**105**	**112**	**120**	**126**	**140**	**138**	**145**	**151**	**159**	**163**	**173**	**189**	**183**	**189**	**198**	**202**
		42	**57**	**72**	**80**	**90**	**105**	**107**	**117**	**126**	**134**	**141**	**150**	**161**	**168**	**173**	**180**	**189**	**199**	**199**	**231**	**223**	**227**	**237**	**244**
22	22	38	48	56	63	70	77	84	91	98	110	110	117	124	130	136	142	148	152	160	163	198	173	182	189
		40	**51**	**62**	**70**	**78**	**84**	**94**	**101**	**108**	**121**	**124**	**130**	**138**	**144**	**150**	**157**	**164**	**169**	**176**	**183**	**198**	**194**	**204**	**209**
		44	**60**	**72**	**83**	**92**	**103**	**112**	**122**	**130**	**143**	**148**	**156**	**164**	**173**	**180**	**187**	**196**	**204**	**212**	**223**	**242**	**237**	**242**	**250**
23	23	38	48	57	65	73	80	89	94	101	108	113	120	127	134	141	146	152	159	164	171	173	207	183	195
		42	**54**	**64**	**72**	**80**	**89**	**98**	**106**	**114**	**119**	**125**	**135**	**142**	**149**	**157**	**163**	**170**	**177**	**184**	**189**	**194**	**230**	**205**	**216**
		46	**63**	**76**	**87**	**97**	**108**	**115**	**126**	**137**	**142**	**149**	**161**	**170**	**179**	**187**	**196**	**204**	**209**	**219**	**227**	**237**	**253**	**249**	**262**
24	24	40	51	60	67	78	84	96	99	106	111	132	125	132	141	152	151	162	164	172	177	182	193	216	204
		44	**57**	**68**	**76**	**90**	**92**	**104**	**111**	**118**	**124**	**144**	**140**	**146**	**156**	**168**	**168**	**180**	**183**	**192**	**198**	**204**	**205**	**240**	**225**
		48	**66**	**80**	**90**	**102**	**112**	**128**	**132**	**140**	**150**	**168**	**166**	**176**	**186**	**200**	**203**	**216**	**218**	**228**	**237**	**242**	**249**	**288**	**262**
25	25	42	54	63	75	78	86	95	101	110	117	120	131	136	145	149	156	162	168	180	182	189	195	204	225
		46	**60**	**68**	**80**	**88**	**97**	**104**	**114**	**125**	**129**	**138**	**145**	**150**	**160**	**167**	**173**	**180**	**187**	**200**	**202**	**209**	**216**	**225**	**250**
		50	**69**	**84**	**95**	**107**	**115**	**125**	**135**	**150**	**154**	**165**	**172**	**182**	**195**	**199**	**207**	**216**	**224**	**235**	**244**	**250**	**262**	**262**	**300**

Anche nel caso di **grandi campioni**, si devono calcolare valori critici differenti se l'ipotesi è a una coda oppure a due code.

Se il **test è a una coda**, secondo la proposta di L. A. **Goodman** del 1954 (vedi ***Kolmogorov-Smirnov tests for psychological research*** in **Psychological Bulletin** Vol. 51, pp. 160-168)
il valore critico viene determinato
mediante

$$\chi^2_{(2)} = 4D^2 \frac{n_1 \cdot n_2}{n_1 + n_2}$$

che ha una distribuzione bene approssimata dal χ^2 **con 2 gradi di libertà**.

Se il **test è a due code**, il valore critico

- alla probabilità $\alpha = 0.05$ è dato da

$$1{,}36 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}$$

- alla probabilità $\alpha = 0.01$ è

$$1{,}63 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}$$

- alla probabilità $\alpha = 0.005$ è

$$1{,}73 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}$$

- alla probabilità $\alpha = 0.001$ è

$$1{,}95 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}$$

Alcuni esempi illustrano la metodologia in modo semplice, ma completo nei suoi passaggi logici.

ESEMPIO 1. (CAMPIONI PICCOLI)

Mediante le cartine al tornasole è possibile misurare il pH di alcuni campioni d'acqua. Metodi analoghi di colorazione vengono usati per confrontare la quantità di fosfati e di nitrati.

Su una scala ordinale con intensità crescente, suddivisa in 8 livelli, sono state riportate le frequenze osservate durante una giornata di rilevazioni in due serie differenti di campioni, raccolti all'ingresso ed all'uscita di un depuratore.

All'ingresso sono stati raccolti 10 campioni e all'uscita 12 campioni, secondo la distribuzione dei valori riportata nella tabella sottostante.

DISTRIBUZIONE OSSERVATA

	Intensità della colorazione							
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**
Ingresso	0	0	0	3	6	0	1	0
Uscita	1	4	4	1	1	1	0	0

Questi dati dimostrano che la quantità di sostanza inquinante contenuta nell'acqua all'uscita del depuratore ha frequenze maggiori di valori bassi? (può dipendere da una media inferiore, da una varianza minore, da variazioni nella simmetria)

Risposta. E' il confronto di 2 piccoli campioni, con un test ad una coda. Infatti in esso si ipotizza che la differenza massima sia nella prima parte della distribuzione, cioè per valori bassi a vantaggio dell'uscita.

1 - Dapprima si trasformano le frequenze assolute in frequenze relative (riga 1 e riga 3)

CALCOLO DELLA DIFFERENZA MASSIMA TRA LE DISTRIBUZIONI CUMULATE

	Intensità della colorazione							
	I	II	III	IV	V	VI	VII	VIII
1) Ingresso freq. Relativa.	0,000	0,000	0,000	0,300	0,600	0,000	0,100	0,000
2) Distribuzione cumulata Ingr.	**0,000**	**0,000**	**0,000**	**0,300**	**0,900**	**0,900**	**1,000**	**1,000**
3) Uscita freq. Relativa	0,084	0,333	0,333	0,084	0,083	0,083	0,000	0,000
4) Distribuzione cumulata Usc.	**0,084**	**0,417**	**0,750**	**0,834**	**0,917**	**1,000**	**1,000**	**1,000**
5) Differenza tra cumulate	0,084	0,417	0,750	0,534	0,017	0,100	0,000	0,000

2) Successivamente si calcolano le cumulate Riga 2 e riga 4);

3) Infine, mediante la serie di differenze tra le due cumulate, si individua lo scarto massimo (riga 5), che risulta uguale a 0,750 (nella classe di colorazione III).

4) Il valore **J** da confrontare con la tabella è

J = 0,750 · 10 · 12 = 90

5) La tabella sinottica che riporta i valori critici per test ad una coda in piccoli campioni,
per $\mathbf{n_1 = 10}$ e $\mathbf{n_2 = 12}$
(anche se è indifferente, perché la tabella dei valori critici è simmetrica)
come valore massimo riporta **74** alla probabilità α **= 0.01**.

6) Il valore calcolato (**J = 90**) è superiore; quindi, la probabilità che la differenza riscontrata sia imputabile al caso è **P< 0.01**. Si rifiuta l'ipotesi nulla, accettando l'ipotesi alternativa che la distribuzione dei dati in entrata e in uscita siano statisticamente differenti.

ESEMPIO 2. (CAMPIONI PICCOLI)

Sovente, all'ambientalista si pone il problema di analizzare la distribuzione territoriale di specie animali o vegetali, per rispondere al quesito se sono più concentrate o più rare in alcune zone oppure se sono distribuite in modo uniforme. Altro quesito importante è se due specie che vivono sullo stesso

territorio hanno distribuzione simile o differente, cioè se occupano le stesse aree con la stessa frequenza, quando esse abbiano un **gradiente di distribuzione**.

Lungo un percorso approssimativamente lineare dalla pianura alla montagna, è stata rilevata la presenza degli individui della specie A e della specie B suddividendo il tragitto in 8 zone consecutive. Si sono raccolte le osservazioni riportate nella tabella sottostante, con il campione della specie A che ha 19 osservazioni e il campione della specie B della quale si sono contati 17 individui.

DISTRIBUZIONE OSSERVATA

	Zone							
	I	II	III	IV	V	VI	VII	VIII
Specie A	4	5	0	2	0	0	1	7
Specie B	1	4	4	6	1	1	0	0

Si può sostenere che le due specie hanno un gradiente di distribuzione differente?

Risposta. E' un test a 2 code e si dispone di piccoli campioni.

1) Dapprima si calcola la differenza massima tra le due cumulate senza considerare il segno, dopo trasformazione delle frequenze assolute in frequenze relative.

CALCOLO DELLA DIFFERENZA MASSIMA TRA LE DISTRIBUZIONI CUMULATE
(in valore assoluto)

	Zone							
	I	II	III	IV	V	VI	VII	VIII
Specie A freq. Relativa	0,211	0,263	0,000	0,105	0,000	0,000	0,053	0,368
Distribuzione cumulata A	**0,211**	**0,474**	**0,474**	**0,579**	**0,579**	**0,579**	**0,632**	**1,000**
Specie B freq. Relativa	0,059	0,235	0,235	0,353	0,059	0,059	0,000	0,000
Distribuzione cumulata B	**0,059**	**0,294**	**0,529**	**0,882**	**0,941**	**1,000**	**1,000**	**1,000**
Differenza tra cumulate	0,152	0,180	0,055	0,303	0,362	0,421	0,368	0,000

2) La differenza massima è **D = 0,421** (nella zona VI) con $\mathbf{n_1 = 19}$ e $\mathbf{n_2 = 17}$.

3) Il valore da confrontare con la tabella dei valori critici

$$\mathbf{J = D \cdot n_1 \cdot n_2 = 0,421 \cdot 19 \cdot 17 = 135,98}$$

è J = 135,98.

4) Per un test a due code

con $n_1 = 19$ e $n_2 = 17$ alla probabilità $\alpha = 0.05$ essa riporta 141.

5) Il valore calcolato è inferiore: la probabilità che le differenze siano imputabili al caso è $P > 0.05$.
Non è possibile rifiutare l'ipotesi nulla.

ESEMPIO 3 (CAMPIONI GRANDI E CONFRONTO TRA TASSI DI SOPRAVVIVENZA).
Con un primo esperimento si è voluto valutare l'effetto di un tossico alla concentrazione del 2%, immettendo in un acquario 150 dafnie per 10 giorni. Con un altro esperimento, si è valutato l'effetto della stessa sostanza alla concentrazione 3% e sono state immesse 200 dafnie.

Nella tabella sottostante, sono riportati i decessi contati ogni giorno nei due differenti esperimenti.

NUMERO OSSERVATO DI DECESSI PER GIORNO

	Giorno I	Giorno II	Giorno III	Giorno IV	Giorno V	oltre V G.
Concentr. 2%	22	43	15	18	16	36
Concentr. 3%	19	39	31	52	59	0

Dai due esperimenti di laboratorio è dimostrato che la concentrazione maggiore abbia una letalità significativamente maggiore, come appare logico attendersi?

Risposta. E' un test ad una coda, con due campioni di grandi dimensioni. Occorre verificare se la concentrazione al 3% ha una frequenza relativa maggiore nei valori bassi e quindi una frequenza relativa minore nei valori alti.

1) Si deve calcolare la differenza massima tra le due distribuzione cumulate osservate, dopo trasformazione nelle frequenze relative.

FREQUENZE RELATIVE DI DECESSI PER GIORNO

	Giorno I	Giorno II	Giorno III	Giorno IV	Giorno V	Oltre V
Concentr. 2%						
Freq. Relativa	0,147	0,287	0,100	0,120	0,106	0,240
Distr. Cumulata	0,147	0,434	0,534	0,654	0,760	1,000
Concentr. 3%						
Freq. Relativa	0,095	0,195	0,155	0,260	0,295	0,000
Distr. Cumulata	0,095	0,290	0,445	0,705	1,000	1,000
Differenza	0,052	0,144	0,089	-0,051	**-0,240**	0,000

2) Per un test ad una coda, la differenza massima nella direzione dell'ipotesi alternativa è **0,240**. La sua significatività è stimata dalla distribuzione χ^2 con 2 gradi di libertà

$$\chi^2_{(2)} = 4 \cdot D^2 \frac{n_1 \cdot n_2}{n_1 + n_2}$$

dove con D = -0,240; $n_1 = 150$; $n_2 = 200$

$$\chi^2_{(2)} = 4 \cdot 0{,}240^2 \frac{150 \cdot 200}{150 + 200} = 0{,}2304 \frac{30000}{350} = 19{,}75$$

si ottiene $\chi^2_{(2)} = 19{,}75$.

3) Alla probabilità $\alpha = 0.001$ il valore critico per 2 gradi di libertà riportato nella tabella sinottica del χ^2 è uguale a 13,82.

4) Il valore calcolato con i dati dell'esempio è superiore: si rifiuta l'ipotesi nulla e si accetta l'ipotesi alternativa di una maggiore letalità del tossico alla concentrazione 3%.

5) Se il confronto fosse stato tra due tossici diversi A e B per un test a due code, in cui verificare le differenze nella distribuzione del numero di cavie decedute per giorno, la stima della significatività della differenza massima (uguale a 0,240) avrebbe dovuto essere confrontata,

- per la probabilità $\alpha = 0.05$, con il valore critico ottenuto dalla relazione

$$1{,}36 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} = 1{,}36 \cdot \sqrt{\frac{150 + 200}{150 \cdot 200}} = 0{,}147$$

che è uguale a 0,147

- e alla probabilità $\alpha = 0.001$ con il valore critico

$$1{,}95 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} = 1{,}95 \cdot \sqrt{\frac{150 + 200}{150 \cdot 200}} = 0{,}211$$

che è uguale a 0,211.

Per entrambe le formule, $n_1 = 150$ e $n_2 = 200$ osservazioni.

La differenza massima riscontrata tra le due distribuzioni cumulate sarebbe risultata significativa.

In un test unilaterale, la differenza è significativa con probabilità $P < 0.001$

Come accennato in precedenza, il test proposto originariamente da Kolmogorov nel 1933 per un campione è stato esteso da Smirnov nel 1939 a due campioni (***On the estimation of the discrepancy between empirical curves of distribution for two independent samples***, pubblicato su **Bull. Moscow Univ. Intern. Ser. (Math**) Vol. 2, pp.3-16), ma sempre per una **scala continua**.

A motivo delle sue applicazioni numerose e importanti, è stato sviluppato anche per gruppi e/o variabili discrete da vari autori. Tra essi,

- W.J. R. **Eplett** nel 1982 (con ***The distributions of Smirnov type two-sample rank tests for discontinuous distributions functions*** pubblicato su **Journal of the Royal Statistical Society** B 44 pp. 361 – 369),
- G. P. **Steck** nel 1969 (con ***The Smirnov two sample tests as rank tests*** , in **Ann. Math. Statist.** Vol. 40, pp. 1449 – 1466)

PER DATI CONTINUI

E' la sua proposta originale. Tra i testi internazionali è riportato in

- **Hollander** Myles, **Wolfe** Douglas A., 1999, ***Nonparametric Statistical Methods***, 2nd ed. John Wiley & Sons, New York, 787 pp.

La metodologia può essere illustrata con un esempio.

1) Si supponga di avere rilevato la quantità di una proteina nel sangue di persone affette (A) da una malattia e del gruppo di Controllo (C), per valutare se le due distribuzioni di dati differiscono. In varie condizioni, la malattia incide sulla media, in altre aumenta sensibilmente la variabilità o la diminuisce oppure modifica la forma della distribuzione, accentuandone l'asimmetria, per la presenza di valori anomali in funzione della gravità della malattia.

A	0,15	1,20	2,30	3,10	3,21	3,58	3,71	4,03	5,22	7,13
C	0,94	0,96	0,99	1,65	1,80	1,99	2,22	2,42	2,48	3,45

Quando due distribuzioni sono differenti, non importa per quale parametro, è logico dedurne che la causa o fattore che le determina siano differenti. Diventa una indicazione importante per individuarli, anche se l'interpretazione e la spiegazione è compito del medico o biologo.

2) Con i valori riportati si costruisce una distribuzione unica: colonna 1 e colonna 2.

(1)	(2)	(3)	(4)	(5)
A	C	Cum. A	Cum. C	$\lvert d \rvert$
0,15		0,1	0,0	0,1
	0,94	0,1	0,1	0,0
	0,96	0,1	0,2	0,1
	0,99	0,1	0,3	0,2
1,20		0,2	0,3	0,1
	1,65	0,2	0,4	0,2
	1,80	0,2	0,5	0,3
	1,99	0,2	0,6	0,4
	2,22	0,2	0,7	0,5
2,30		0,3	0,7	0,4
	2,42	0,3	0,8	0,5
	2,48	0,3	0,9	**0,6**
3,10		0,4	0,9	0,5
3,21		0,5	0,9	0,4
	3,45	0,5	1,0	0,5
3,58		0,6	1,0	0,4
3,71		0,7	1,0	0,3
4,03		0,8	1,0	0,2
5,22		0,9	1,0	0,1
7,13		1,0	1,0	0,0

3) Per ognuna delle due distribuzioni si calcola la cumulata fino a quel punto, con le frequenze relative (Per motivi didattici è stata scelta una forma molto semplice; quindi i due campioni hanno 10 dati ognuno).
La colonna 3 rappresenta la cumulata fino a quel dato dei valori continui riportati nella colonna 1.
La colonna 4 rappresenta la cumulata fino a quel dato dei valori continui riportati nella colonna 2.

4) Si calcolano tutte le differenze tra le due cumulate, come nella colonna 5.
La differenza massima è in coincidenza dal valore 2,48 e $d_{max} = 0{,}6$

5) Per i valori critici è possibile utilizzare la tabella riportata (anche se esistono proposte molto più sofisticate) A questo scopo si calcola J

$$J = D \cdot n_1 \cdot n_2 = 0{,}6 \cdot 10 \cdot 10 = 60$$

ottenendo J = 135,98.

- Nella tabella, per un test bilaterale, con $\mathbf{n_1 = 10}$ e $\mathbf{n_2 = 10}$ sono riportati
- alla probabilità $\alpha = 0.10$ quando $J \geq 60$,
- alla probabilità $\alpha = 0.05$ quando $J \geq 70$,
- alla probabilità $\alpha = 0.01$ quando $J \geq 80$.

Di conseguenza, il test non risulta significativo, avendo una probabilità P = 0.10

6) Utilizzando la formula per grandi campioni, il valore critico alla probabilità $\alpha = 0.05$ è ottenuto dalla relazione

$$1{,}36 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} = 1{,}36 \cdot \sqrt{\frac{10+10}{10 \cdot 10}} = 1{,}36 \cdot \sqrt{0{,}2} = 0{,}608$$

e risulta uguale a 0,608.
Poiché la differenza massima (0,6) è inferiore al valore critico calcolato, con P > 0.05 non è possibile rifiutare l'ipotesi nulla. Ma si potrebbe affermare che è molto vicino al valore critico. In realtà la situazione è differente.

A**vvertenza importante.** Con la probabilità stimata leggermente superiore a 0.05, si sarebbe potuto affermare che la risposta è tendenzialmente significativa. Purtroppo è ottenuta con la **formula approssimata** che, come spesso avviene, determina probabilità inferiori alla realtà. Quindi con essa aumenta la probabilità di trovare una differenza, quando nella realtà essa non esiste. Tra le due

possibilità d'errore (probabilità minore o maggiore del reale) per lo statistico quella è più grave, che non dovrebbe mai commettere, in quanto induce ad una scelta errata e afferma una cosa non corretta.
Usando i computer, che possono avere in memoria tabelle molto complesse, la probabilità fornita dovrebbe non avere questi errori di **approssimazione asintotica**.