

# La significatività assoluta e condizionata, considerazioni sulla logica inferenziale e sulla rappresentatività campionaria, nel test sulla differenza tra proporzioni per piccoli campioni

Francesco Gori, Gianfranco Cavedon  
francesco.gori@uniroma1.it  
francesgori@hotmail.it  
Dipartimento di Scienze Statistiche, "Sapienza" Università di Roma

## Abstract

In this paper we use the test on the difference between proportions, in discrete field, to trait the conditioned and absolute significance and to give some reflection about a logic reversal of the hypothesis testing. In order to discuss sample representativeness an experimental model is formulated.

The work aims to propose a test to asses whether two samples with respective proportions  $p_1$  and  $p_2$  come from the same dichotomous population with given parameter  $\pi$  (state of Nature), assessed or proposed by the researcher. The discussion of the subsequent topics is focused on  $\pi$ .

The problem is formulated as hypothesis test, then we discuss the test statistic and some critical remarks on the Normal approximation, based on skewness index and kurtosis. The procedure gives the sample distributions under binomial and hypergeometric distributions. Tables are presented to illustrate the variability of the critical values when  $\pi$  and the sample size vary. The procedure has been implemented into a software (DPT.exe) that can be requested to and freely provided by the Author.

This work has an operational function because it uses the application DPT and given the detailed processing phases carried out and described by the software, it is also a valuable didactical instrument.

The work brings together the contents of a previous Rapporto Tecnico n. 11, Maggio 2007, DSPSA, La Sapienza, Roma, and some additional considerations relating to Sections 3.2, 4.1 and 4.2 ready in June 2008, and that was not possible to publish before.

## RIASSUNTO

In questo lavoro si utilizza il test sulla differenza tra proporzioni, nel discreto, per trattare la significatività assoluta e condizionata e per una riflessione su un'inversione logica della verifica delle ipotesi. Nell'ambito di considerazioni sulla rappresentatività campionaria viene proposto un modello sperimentale.

Viene trattata la verifica delle ipotesi circa la provenienza o meno di due campioni che hanno prodotto le due proporzioni  $p_1$  e  $p_2$ , dalla stessa popolazione dicotomica, caratterizzata dal parametro  $\pi$  (stato di Natura), valutato o proposto dal ricercatore. Sullo stato di Natura  $\pi$  è incentrata anche la trattazione degli argomenti successivi.

Viene presentato il problema in termini di sistema di ipotesi, la variabile utilizzata ed osservazioni critiche in merito all'approssimazione mediante la Normale, basate sull'indice di asimmetria e di curtosi. Il procedimento costruisce le distribuzioni campionarie utilizzando le distribuzioni binomiale ed ipergeometrica. Sono riportate Tabelle per una esemplificazione della variabilità dei valori critici al variare di  $\pi$  e delle numerosità campionarie.

Il procedimento è stato programmato. Il programma DPT.exe può essere fornito gratuitamente. L'interesse del lavoro è operativo data la presenza del programma, che, dato il dettaglio delle fasi di elaborazione effettuate e descritte, è utile per la didattica.

Il lavoro riunisce il contenuto di un precedente Rapporto Tecnico n. 11, Maggio 2007, DSPSA, La Sapienza, Roma, ed alcune considerazioni aggiuntive relative alle Sezioni 3.2, 4.1 e 4.2, già pronte nel Giugno 2008, e che non è stato possibile pubblicare prima.

**Keywords:** absolute and conditioned significance, difference, proportions, small samples, sample representativeness

## 1. PRESENTAZIONE E ASPETTI PARTICOLARI DEL LAVORO

In questo lavoro la differenza tra proporzioni viene considerata come un caso di differenza tra medie [10]. Trattando piccoli campioni, il procedimento utilizza le distribuzioni campionarie discrete, binomiale ed ipergeometrica, dato  $\pi$ . Il test quindi non fa riferimento alla distribuzione Normale o al chi quadro, non introduce le correzioni dei vari metodi esistenti in letteratura e, a differenza del test di Fisher<sup>1)</sup>, è libero dal vincolo ai totali marginali del numero dei successi e degli insuccessi [1], [2], [16].

Il sistema di ipotesi è,  $H_0: \pi_1 = \pi_2 = \pi$ , per  $\pi = .05, .01, \dots, .90, .95$  ( $0 < \pi < 1$ ) contro  $H_1: \pi_1 \neq \pi_2$  mono o bidirezionale. Altri Autori considerano test, incondizionati (non vincolati) e basati sull'intervallo di confidenza: o di  $\pi$  (si rifiuta  $H_0$  se al suo interno gli esiti portano tutti al rifiuto) [1]; o di  $\pi_1 - \pi_2$  (si rifiuta  $H_0$  se i limiti, ottenuti tramite la differenza sperimentale, non comprendono lo 0) [4], [5]. In letteratura  $\pi$  è ritenuto un parametro di disturbo. In questo lavoro lo stato di Natura  $\pi$  è fissato dal ricercatore ed ha un ruolo fondamentale per la trattazione dei vari argomenti. In precedenza [14] era stata trattata una procedura riferita alle distribuzioni ottenute mediante estrazioni con e senza reimmissione. Quel lavoro considerava gruppi della stessa numerosità  $n_1 = n_2 = n$  ( $3 \leq n \leq 30$ ), preferibile nel caso di piccoli campioni in quanto gruppi di diversa ampiezza non hanno lo stesso contenuto informativo, ma soprattutto perché limitava ad una sola tabella ( $28 \times 10$ ) il prospetto dei valori critici (Appendice 4). In questa sede, viene trattato anche il caso di due campioni aventi ampiezza diversa. Il lavoro utilizza un programma, DPT (Sezioni 5.1 e 5.2), procedura analitico-numerica che, al variare della forma della popolazione ( $\pi$ ), determina i valori critici dato  $\alpha$ , e/o confronta il p-value sulle code della distribuzione discreta con quello ottenuto dalla Normale. Le popolazioni sono costruite dall'operatore, orientato, per lo stato di Natura  $\pi$  sotto  $H_0$ , da posizioni soggettive o da elementi (stima e intervallo di confidenza) forniti dal programma.

L'approccio mediante piccoli campioni è utile nel caso di popolazioni limitate, come ad esempio nel campo della psicologia, ricerche di laboratorio o nei casi di fenomeni difficilmente rilevabili.

### 2.1 ESTRAZIONE CON REIMMISSIONE, VARIABILE UTILIZZATA, CARATTERISTICHE DELLA DISTRIBUZIONE CAMPIONARIA

Dati due campioni indipendenti di numerosità  $n_1$  ed  $n_2$ , estratti sotto  $H_0$  da una popolazione a due alternative A e B con probabilità  $\pi$  e  $1 - \pi$ , vengono osservati rispettivamente  $a_1$  e  $a_2$  successi. La differenza tra proporzioni generalmente viene trattata mediante la variabile:

$$(p_1 - p_2) / [pq(1/n_1 + 1/n_2)]^{1/2} \quad (1)$$

in cui  $p_1 = a_1/n_1$  e  $p_2 = a_2/n_2$  sono le proporzioni prodotte dai due campioni, mentre la

$$p = (n_1 p_1 + n_2 p_2) / (n_1 + n_2) = (a_1 + a_2) / (n_1 + n_2), \quad (2)$$

e la  $q = 1 - p$ .

La variabile (1) utilizza lo stimatore<sup>2)</sup> di massima verosimiglianza (2) per valutare lo stato di Natura  $\pi$  (sotto  $H_0$ ), ignoto, e fa riferimento alla Normale standardizzata; procedura usata spesso, erroneamente, per numerosità limitate. Nel caso attuale si utilizzano: la variabile differenza tra proporzioni  $dp = p_1 - p_2$  e, in un primo approccio, la stima data dalla (2), ma non per standardizzare la  $dp$ , bensì per valutare  $\pi$ , parametro della popolazione da cui provengono i due campioni sotto  $H_0$ , al fine di costruire la distribuzione esatta [19] della  $dp$ . Tale distribuzione, analogamente al caso del test binomiale quando il campione ha numerosità  $\leq 30$ , risulterà preferibile all'uso dell'approssimazione mediante la Normale.

Dato  $\pi$  sotto  $H_0$ , faremo riferimento all'estrazione con reimmissione in cui la frequenza:

$$fr[(p_1 - p_2) | \pi] = fr(p_1 | \pi) fr(p_2 | \pi) = fr(n_1, n_2, a_1, a_2 | \pi) = \binom{n_1}{a_1} \pi^{a_1} (1 - \pi)^{n_1 - a_1} \times \binom{n_2}{a_2} \pi^{a_2} (1 - \pi)^{n_2 - a_2} \quad (3)$$

La distribuzione si ottiene considerando tutte le possibili differenze  $dp = p_1 - p_2 = a_1/n_1 - a_2/n_2$  ( $a_1 = 0, 1, \dots, n_1$ ;  $a_2 = 0, 1, \dots, n_2$ ) e raggruppando le frequenze (3), a parità di  $dp$ , crescente, variabile da  $-1$  ( $0/n_1 - n_2/n_2$ ) a

$1$  ( $n_1/n_1 - 0/n_2$ ). Dati  $n_1$  ed  $n_2$  i successivi valori discreti della variabile  $dp$  sono sempre gli stessi, per qualsiasi valore di  $\pi$ . La  $f(dp | \pi) = \sum_{\forall (p_1 - p_2) = dp} fr[(p_1 - p_2) | \pi]$  è la valutazione della probabilità di quel valore  $dp$ .

Sotto  $H_0$  ( $\pi_1 = \pi_2 = \pi$ ) la distribuzione ha media:  $\mu_1(dp) = E(dp) = E(p_1 - p_2) = E(p_1) - E(p_2) = \pi - \pi = 0$ , e varianza:  $\mu_2'(dp) = E[(dp-0)^2] = E(dp^2) = \pi(1-\pi)(1/n_1 + 1/n_2) = \sigma_{dp}^2$ .

Per segnalare gli aspetti che sconsigliano l'uso dell'approssimazione mediante la Normale, una prima osservazione è che se nella popolazione sotto  $H_0$ ,  $\pi \neq 1-\pi$ , la distribuzione della variabile  $dp$  non è simmetrica, analogamente a quanto accade nel caso di un solo campione. Nel caso attuale infatti  $\mu_3(dp) = E(dp^3) = E[(p_1 - p_2)^3]$ , e poiché i due campioni sono indipendenti, nel generico termine dello sviluppo si ha:  $E[(p_1^l p_2^m)] = E(p_1^l) E(p_2^m)$ ; applicando note espressioni [11] in definitiva si ottiene dopo alcuni passaggi il momento terzo centrale:

$$\mu_3'(dp) = \pi(1-\pi)(1-2\pi)(n_2^2 - n_1^2)/(n_1^2 \times n_2^2)$$

quindi l'indice di asimmetria:

$$\gamma_1(dp) = \frac{\mu_3'(dp)}{\sigma_{dp}^3} = \frac{(1-2\pi)(n_2 - n_1)}{[\pi(1-\pi)n_1n_2(n_1 + n_2)]^{1/2}}; \quad (4)$$

$\gamma_1$  è nullo, come per la Normale, per  $\pi=1-\pi$ , ma raramente la (2) porta a valori  $p = q$ . La situazione migliora se  $n_1 = n_2 = n$ , come risulta anche nel precedente lavoro [14], in cui la distribuzione è sempre simmetrica.

Una seconda osservazione si può fare considerando la curtosi. Infatti, l'espressione del momento quarto centrale, ottenuto con procedimento analogo a quanto visto per il terzo, è:

$$\mu_4'(dp) = \frac{\pi(1-\pi)}{n_1^2 n_2^2} \left\{ 3(n_1+n_2)^2 \pi(1-\pi) + \frac{n_1^3 + n_2^3}{n_1 n_2} [1-6\pi(1-\pi)] \right\}$$

e si ha dopo alcuni passaggi l'indice di curtosi:

$$\gamma_2(dp) = \frac{\mu_4'(dp)}{\sigma_{dp}^4} - 3 = \frac{(n_1^3 + n_2^3)[1-6\pi(1-\pi)]}{n_1 n_2 (n_1 + n_2)^2 \pi(1-\pi)} \quad (5)$$

che risulta = 0, come per la Normale, se  $\pi$  (o  $1-\pi$ ) =  $.5 \pm \sqrt{3}/6$ . Quindi, facendo riferimento alla curtosi, per piccoli campioni la Normale approssima in maniera soddisfacente una distribuzione binomiale se in questa  $\pi$  (o  $1-\pi$ ) ha valore .7886751. Ma tale valore di  $\pi$  che annulla la curtosi fornisce per l'asimmetria l'espressione:  $\mp (n_2 - n_1) / [n_1 n_2 (n_1 + n_2) / 2]^{1/2}$  che si annulla solo per  $n_1 = n_2$ . L'approssimazione mediante la Normale migliora al crescere delle numerosità  $n_1$  e  $n_2$ <sup>3)</sup>. Il parametro  $\pi$  e le numerosità  $n_1$  ed  $n_2$  determinano la forma della distribuzione campionaria della variabile  $dp$ .

Se  $\pi > 1-\pi$  o se  $\pi < 1-\pi$ , dalla (4) le distribuzioni risultano con asimmetria invertita di segno, e ciò dipende dal fatto che nella espressione (3)  $\pi$  e  $1-\pi$  giocano un ruolo simmetrico e scambiabile, legato alle  $a_1$  e  $a_2$  ed alle rispettive numerosità  $n_1$  ed  $n_2$ , quindi anche l'andamento delle frequenze nelle code risulta scambiato. In conseguenza i valori critici possono risultare scambiati in valore assoluto e segno.

Mediante il programma DPT di cui parleremo nelle Sezioni 5.1 e 5.2 sono state costruite le Tabelle che seguono.

## 2.2 TABELLE-ESEMPIO DI VALORI CRITICI DEL TEST, AL VARIARE DI $\pi$ E DELLE NUMEROSITÀ CAMPIONARIE

Per avere un'altra indicazione di come i valori di  $\pi$ ,  $n_1$  ed  $n_2$  influiscano sulla forma delle distribuzioni, sui valori critici<sup>4)</sup> e quindi sul risultato del test, sono state costruite le Tabelle 1 e 2 per  $\alpha = .025$  ed  $\alpha = .05$ , al variare di  $\pi$ . I valori critici, inferiore  $dp'$  e superiore  $dp''$  dipendono da  $\pi$ ,  $n_1$ ,  $n_2$  ed  $\alpha$ .

Le Tabelle sono state limitate per motivi di spazio solo al caso di numerosità  $n_1$  ed  $n_2$  non troppo differenti, per quanto detto sopra, come nel seguente prospetto; per numerosità uguali o differenti ( $3 \leq n \leq 30$ ) i valori critici possono essere trovati mediante programma:

$$\begin{array}{lll} n_1 = 5 \text{ ed } n_2 = 10; & n_1 = 10 \text{ ed } n_2 = 15; & n_1 = 15 \text{ ed } n_2 = 20; \\ n_1 = 20 \text{ ed } n_2 = 25; & n_1 = 20 \text{ ed } n_2 = 30; & n_1 = 25 \text{ ed } n_2 = 30; \end{array}$$

per ipotesi  $H_0: \pi_1 = \pi_2 = \pi$ , essendo  $\pi = .05$  (.95), .1 (.90), .15 (.85), .2 (.8), .25 (.75), ..., .5.

Le Tabelle hanno un limitato valore applicativo, perché esigono che le numerosità campionarie, differenti<sup>5)</sup>, abbiano i valori ivi riportati; ma vanno viste soprattutto nel loro contributo illustrativo della variabilità dei valori critici  $dp'$  e  $dp''$ .



$a_2=16$ ,  $dp_s = -.6286$ ,  $p=.5$ , il p-value è 0.00008 (discr.) e 0.0001 (Nor.); ma le ultime sono situazioni che riguardano gli estremi delle distribuzioni. E ancora, ad es. per  $n_1=14$ ,  $n_2=22$ , dati i risultati  $a_1=5$ ,  $a_2=16$  quindi  $dp_s = -.37013$ , con valutazioni di  $\pi : .1, .2, \dots, .9$ , per la Normale notiamo la simmetria nei risultati sia di  $z$  che del p-value:

$z$	-3.60875	-2.70656	-2.36248	-2.20990	-2.16530	-2.20990	-2.36248	-2.70656	-3.60875
p-value. Nor.	.00015	.00340	.00908	.01356	.01518	.01356	.00908	.00340	.00015
p-value discr.	.00004	.00193	.00749	.01316	.01651	.01568	.01006	.00440	.00006

e, come negli esempi precedenti, vediamo che i p-value sono superiori a quelli del discreto solo negli estremi di alcune distribuzioni e con valori trascurabili rispetto a .025 e .05; più in generale, a parità di stima di  $\pi$ , per tali valori di  $\alpha$ , utilizzando la distribuzione discreta anziché la Normale, il test sembrerebbe più prudentiale.

Nel seguito, per gli argomenti che verranno trattati considereremo la distribuzione discreta.

### 2.3 CENNI SULLA POTENZA DEL TEST

Come noto, la potenza  $1 - \beta$  di un test dipende dal valore della distanza tra il parametro sotto l'ipotesi nulla e il parametro sotto l'ipotesi alternativa, e dalla numerosità campionaria.

Nel nostro caso viene calcolata utilizzando la distribuzione campionaria sotto  $H_1: \pi_1 \neq \pi_2$ . Circa la distanza tra l'ipotesi nulla e l'ipotesi alternativa faremo in seguito alcune considerazioni; prima trattiamo la distribuzione sotto l'ipotesi alternativa. Dati i valori dei parametri  $\pi_1$  e  $\pi_2$ , le numerosità  $n_1$  ed  $n_2$  ed i numeri dei successi  $a_1$  e  $a_2$ , la frequenza  $f(p_1 - p_2)$  dell'evento  $p_1 - p_2$  per gruppi indipendenti nel discontinuo ed estrazioni con reimmissione è:

$$f(n_1, n_2, a_1, a_2 | \pi_1, \pi_2) = \binom{n_1}{a_1} \pi_1^{a_1} (1 - \pi_1)^{n_1 - a_1} \times \binom{n_2}{a_2} \pi_2^{a_2} (1 - \pi_2)^{n_2 - a_2} \quad (6)$$

in cui  $a_1$  può assumere i valori:  $0, 1, \dots, n_1$ ; e  $a_2$  i valori:  $0, 1, \dots, n_2$ .

La distribuzione della variabile  $dp = p_1 - p_2 = a_1/n_1 - a_2/n_2$ ,  $f(dp | \pi_1, \pi_2)$ , si ottiene come nel caso della (3).

La distribuzione ha media  $E(p_1 - p_2) = E(p_1) - E(p_2) = \pi_1 - \pi_2$ .

La varianza, nell'estrazione con reimmissione è:  $\sigma_{dp}^2 = \pi_1(1 - \pi_1)/n_1 + \pi_2(1 - \pi_2)/n_2$ .

Parliamo ora della distanza tra le due ipotesi. Date  $H_0: \pi_1 = \pi_2$  e  $H_1: \pi_1 > \pi_2$ , oppure  $\pi_1 < \pi_2$ , osserviamo che sotto  $H_0$  la distribuzione della variabile  $dp$  ha media  $\pi_1 - \pi_2 = 0$ ;

per un test monolaterale destro sotto  $H_1$ , la media della distribuzione è  $\pi_1 - \pi_2 > 0$ ;

per un test monolaterale sinistro sotto  $H_1$ , la media è  $\pi_1 - \pi_2 < 0$ ;

per un test bilaterale  $H_1: \pi_1 \neq \pi_2$ , la media può essere  $\pi_1 - \pi_2 > 0$ , oppure  $\pi_1 - \pi_2 < 0$ .

Nel nostro caso la distanza è data dalla differenza tra la media  $\pi_1 - \pi_2 \neq 0$ , relativa all'ipotesi alternativa, e la media 0 relativa all'ipotesi nulla, e quindi è data da  $\pi_1 - \pi_2 - 0 = \pi_1 - \pi_2 = \delta$ , differenza tra i due parametri.

Ma una determinata differenza può essere realizzata con diversi valori dei parametri  $\pi_1$  e  $\pi_2$ , parametri che figurano nella (6) e che quindi possono dar luogo a diversi valori di  $1 - \beta$ . Per distinguere le varie situazioni si farà riferimento a  $\delta$  e allo stato di Natura  $\pi_1$ . Una funzione di distribuzione sotto  $H_1$  sarà indicata con  $f(dp | \delta, \pi_1)$  e, per  $\delta \neq 0$ , farà parte dell'insieme  $f(dp | H_1)$ . Per  $\delta = 0$  vale l'ipotesi  $H_0: \pi_1 = \pi_2$ .

Consideriamo un test monolaterale destro. Dato  $\alpha$ , si determina il valore critico  $dp^*$  incluso<sup>6)</sup> nella zona di rifiuto sulla coda destra della distribuzione sotto  $H_0$ . Quindi, fissati i parametri  $\pi_1$  e  $\pi_2$  ( $\pi_1 > \pi_2$ ), mediante la  $f(dp | \delta, \pi_1)$  si determina la distribuzione sotto  $H_1$ , la cui media è posizionata a destra della media della distribuzione sotto  $H_0$ . Disponendo della distribuzione sotto  $H_1$ , si calcola la frequenza degli errori di seconda specie  $\beta$ , per valori della variabile  $dp$ , da  $dp = -1$ , fino al valore  $dp^*$ , escluso, tenendo presente che il valore critico è incluso nella zona di rifiuto:

$$\beta = \sum_{dp=-1}^{< dp^*} f(dp | \delta, \pi_1) \quad (7)$$

Per differenza si ottiene la potenza del test,  $1 - \beta$  <sup>7)</sup>.

Per un test monolaterale sinistro,  $\pi_1 - \pi_2 < 0$ , la distribuzione sotto  $H_1$  si trova alla sinistra della distribuzione sotto  $H_0$ . In tal caso, incluso  $dp'$ , la potenza del test è data da:

$$1 - \beta = \sum_{dp=-1}^{dp'} f(dp | \delta, \pi_1). \quad (8)$$

Ciò che interpreta la direzionalità del test è l'espressione usata per calcolare la potenza con riferimento agli estremi delle sommatore.

Per un test bilaterale possiamo considerare l'area  $\beta$  sempre contenuta tra  $dp'$  e  $dp''$ , esclusi:

$$\beta = \sum_{dp \Rightarrow dp'}^{< dp''} f(dp | \delta, \pi_1), \quad (9)$$

da cui come sopra si ottiene  $1 - \beta$ . In questo e negli altri due casi si tratta poi di vedere cosa accade della potenza quando l'ipotesi alternativa è posta più verso sinistra o più verso destra. Ma, come già accennato, il problema nasce dal fatto che una stessa differenza,  $\pi_1 - \pi_2$ , si può ottenere con valori diversi di  $\pi_1$ .

Fissati  $n_1$  ed  $n_2$ , per ogni valore di  $\pi_1$  (.05, .1, .15, .2, ... .95), con le formule viste si determina l'andamento della potenza al variare della differenza  $\delta$ . Per facilitare la rappresentazione grafica per ogni valore di  $\pi_1$  conviene far variare  $\pi_2$  da .95 a .05 in modo che  $\delta$  possa assumere valori crescenti, da -.9 (per  $\pi_1 = .05$ ) a .9 (per  $\pi_1 = .95$ ):

$\pi_1$	$\pi_2$	Prospetto delle differenze $\delta = \pi_1 - \pi_2$															
.05	da .95 a .05,	-9	-.85	-.8	-.75	-.7	-.65	-.6	-.55	.....	-.3	-.25	-.2	-.15	-.1	-.05	0
.1	“ “ “ “	-.85	-.8	-.75	-.7	-.65	-.6	-.55	-.5	.....	-.25	-.2	-.15	-.1	-.05	0	.05
.15	“ “ “ “	-.8	-.75	-.7	-.65	-.6	-.55	-.5	-.45	.....	-.2	-.15	-.1	-.05	0	.05	.1
.2	“ “ “ “	-.75	-.7	-.65	-.6	-.55	-.5	-.45	-.4	.....	-.15	-.1	-.05	0	.05	.1	.15
.....																	
.5	“ “ “ “	-.45	-.4	-.35	-.3	-.25	-.2	-.15	-.1	.....	.15	.2	.25	.3	.35	.4	.45
.....																	
.85	“ “ “ “	-.1	-.05	0	.05	.1	.15	.2	.25	.....	.5	.55	.6	.65	.7	.75	.8
.9	“ “ “ “	-.05	0	.05	.1	.15	.2	.25	.3	.....	.55	.6	.65	.7	.75	.8	.85
.95	“ “ “ “	0	.05	.1	.15	.2	.25	.3	.35	.....	.5	.65	.7	.75	.8	.85	.9

Dal prospetto notiamo che gli estremi del campo di variazione delle differenze ( $-9 \leftrightarrow 0$ ;  $-.85 \leftrightarrow .05$ ;  $-.8 \leftrightarrow .1$ ; ecc.) variano con  $\pi_1$ , e soprattutto che, seguendo il verso della diagonale secondaria,  $\delta$  assume uguali valori per valori di  $\pi_1$  diversi, quindi è opportuno indicare i valori delle differenze con  $\delta | \pi_1$ . Osserviamo che lo stato di Natura  $\pi_1$  è l'elemento aggiuntivo rispetto alla trattazione (nota) che considera  $\delta$  e le ampiezze campionarie.

Per passare da un valore di  $\delta | \pi_1$  ai valori dei parametri utilizzabili nella (6), partendo da  $\delta | \pi_1 = \pi_1 - \pi_2$  si ha  $\pi_2 = \pi_1 - \delta | \pi_1$ . Ad esempio, terza riga e seconda colonna, partendo da  $\delta | \pi_1 = -.75 | .15$ , i valori dei parametri sono  $\pi_1 = .15$  e  $\pi_2 = .15 - (-.75) = .90$ ; nella penultima riga e prima colonna,  $\delta | \pi_1 = -.05 | .9$  significa  $\pi_1 = .9$  e  $\pi_2 = .9 - (-.05) = .95$ , mentre  $\delta | \pi_1 = .8 | .9$  significa  $\pi_1 = .9$  e  $\pi_2 = .9 - .8 = .1$ .

Fissati  $n_1, n_2, \pi_1$  e  $\delta | \pi_1$ , per calcolare la potenza, prima si costruisce (da programma) la distribuzione sotto  $H_0$  partendo dalla (3) con parametro  $\pi_1$ , e dato  $\alpha$  si determinano i valori critici  $dp'$  e/o  $dp''$  (come in Tabella 1 e 2); poi, facendo riferimento alla (6), si costruisce la distribuzione sotto  $H_1$ , e in questa, con le formule sopra descritte, si calcola  $1 - \beta$ . L'operazione va ripetuta al variare di  $\pi_1$ , di  $\delta | \pi_1$ , e andrebbe fatta anche per  $n_1$  ed  $n_2$ , crescenti. Dal punto di vista operativo, nel calcolo delle somme che danno  $1 - \beta$ , per evitare errori dovuti ad arrotondamenti, nella distribuzione sotto  $H_1$  si è utilizzato il numero d'ordine (n.o.) che un valore critico occupa nella distribuzione sotto  $H_0$  (dati  $n_1$  ed  $n_2$ , i successivi valori della variabile  $dp$  nelle distribuzioni sotto  $H_0$  e  $H_1$  sono gli stessi).

Facciamo alcuni esempi di calcolo della potenza considerando un test monolaterale destro, un test monolaterale sinistro e un test bilaterale, riportati in Appendice 1, 2 e 3. Per ogni esempio usiamo tre valori di  $\pi_1$ : il primo .15,

il secondo (intermedio) .5, il terzo .85. I valori .15 e .85 sono stati scelti volutamente simmetrici rispetto a .5 per mostrare eventuali simmetrie nei risultati. Le ampiezze sono  $n_1 = 5$  ed  $n_2 = 10$ .

Test monolaterale **destro**, in **Appendice 1**, per  $\pi_1 = .15$ , riportiamo per intero i valori di  $\delta |.15$ :

-0.8, -0.75, -0.7, -0.65, -0.6, -0.55, -0.5, -0.45, -0.4, -0.35, -0.3, -0.25, -0.2, -0.15, -0.1, -0.05, 0, .05, .1

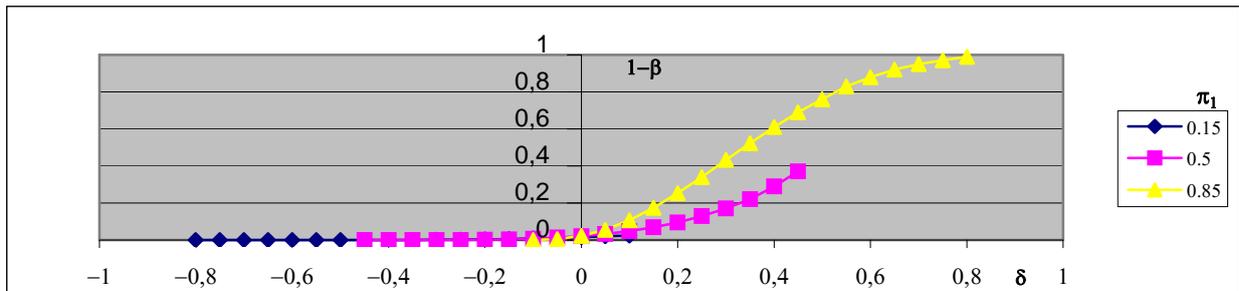
Come già detto, per  $\pi_1 = .15$  e  $\pi_2 = .15$  costruiamo la distribuzione sotto  $H_0$  e, fissato  $\alpha = 0.025$ , determiniamo il valore critico  $dp'' = .5$  che ha n.o. 16. Quindi, con riferimento ai valori sopra visti di  $\delta |.15$ , costruiamo le distribuzioni sotto  $H_1$  e con quel n.o., usando la (7) che ci dà  $\beta$ , calcoliamo le potenze. I risultati sono nel Prospetto 1. Al crescere di  $\delta |.15$  la potenza risulta trascurabile ma crescente, con andamento a J e arriva a 0.0245.

Per  $\pi_1 = .5$  il valore critico sotto  $H_0$  è  $dp'' = .6$ , con n.o. 17. Nel Prospetto 2, al crescere di  $\delta |.5$  da -.45 a .45 (campo di variazione più a destra del precedente) la potenza è crescente con andamento a J e arriva a 0.3728.

Per  $\pi_1 = .85$   $dp'' = .400$  con n.o. 15. Nel Prospetto 3, al crescere di  $\delta |.85$  con valori che vanno da -.1 a .8, i risultati confermano l'andamento a J e la potenza arriva a 0.9864.

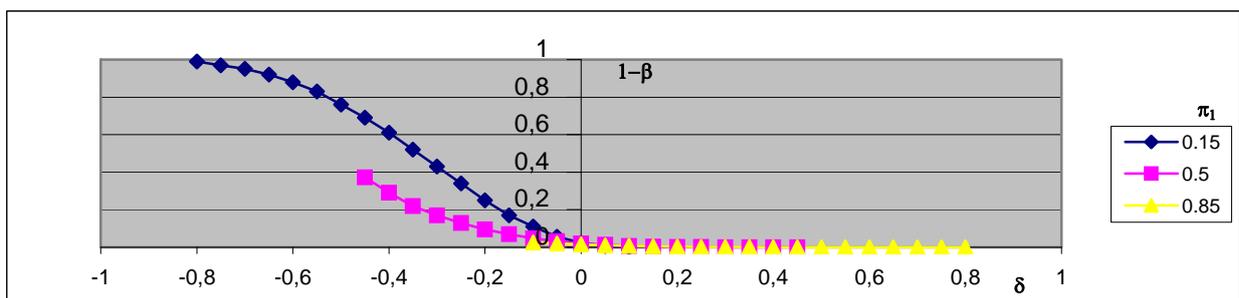
Osserviamo che: a) per  $\delta = 0$  le potenze: 0.0154, 0.0206 e 0.0227 coincidono con le frequenze nel discreto  $\leq \alpha$  delle zone di rifiuto di  $H_0$ ; b) al crescere di  $\delta$  la potenza è crescente, con curve a J che slittano verso destra (al crescere di  $\pi_1$ ) e hanno ordinata massima crescente. In un grafico per  $\pi_1 = .15$ ;  $\pi_1 = .5$ ;  $\pi_1 = .85$ , e  $\delta$  crescente, abbiamo i tre andamenti della potenza.

Test monolaterale destro, grafico della potenza per  $n_1 = 5$ ,  $n_2 = 10$ ,  $\alpha = 0.025$ , al variare di  $\pi_1$  e  $\delta$



Test monolaterale **sinistro**, in **Appendice 2**, al crescere di  $\delta$  la potenza è decrescente, con valori massimi decrescenti: 0.9864, 0.3728 e 0.0245; graficamente le curve, ad L, slittano verso destra al crescere di  $\pi_1$ ; risultato specularmente simile al precedente per l'inversione delle code delle distribuzioni e la simmetria dei valori di  $\pi_1$ .

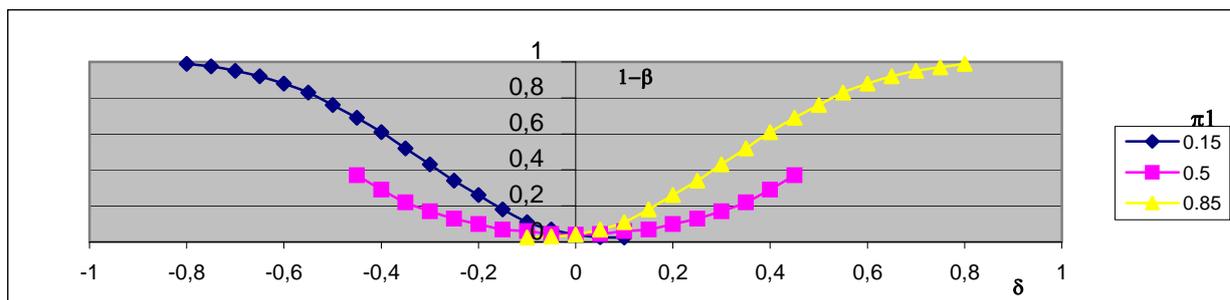
Test monolaterale sinistro, grafico della potenza per  $n_1 = 5$ ,  $n_2 = 10$ ,  $\alpha = 0.025$ , al variare di  $\pi_1$  e  $\delta$



Test **bilaterale**, in **Appendice 3**, con riferimento ai valori crescenti di  $\delta$  notiamo una simmetria inversa tra i valori della potenza per  $\pi_1 = .15$  e  $\pi_1 = .85$ . Più in dettaglio, per  $\pi_1 = .15$ , Prospetto 1, da un massimo di 0.9864 la potenza decresce con andamento a L. Per  $\pi_1 = .85$ , Prospetto 3, la potenza cresce con andamento a J fino a 0.9864.

Nel caso intermedio  $\pi_1 = .5$ , Prospetto 2, notiamo una simmetria nei valori della potenza, decrescenti da 0.3728 fino a 0.0413 ( $H_0$ ), poi crescenti fino a 0.3728, graficamente con andamento a V, unico caso di curva simmetrica.

Test bilaterale, grafico della potenza per  $n_1 = 5, n_2 = 10, \alpha = 0.05$ , al variare di  $\pi_1$  e  $\delta$



Conclusione. Dai Prospetti notiamo l'ordinata in neretto per  $\delta = 0$ , frequenza degli errori di prima specie nel discreto  $\leq \alpha$  ( $H_0$ ). Inoltre notiamo per  $\pi_1 = .15$  e  $\delta = 0$ : nel test monolaterale destro una potenza pari a 0.0154, nel sinistro pari a 0.0227, la loro somma dà 0.0381 (del bilaterale). Per  $\pi_1 = .5$  e  $\delta = 0$  nei test monolaterali esemplificati la potenza è 0.0206 mentre nel test bilaterale è il doppio, 0.0413. Per  $\pi_1 = .85$  e  $\delta = 0$  nel test monolaterale destro si ha 0.0227, nel monolaterale sinistro 0.0154, la loro somma dà 0.0381 (bilaterale).

I risultati esposti riguardano la potenza nel piano  $\delta, 1 - \beta$ , dato lo stato di Natura  $\pi_1$ ; i successivi valori di  $\pi_1$  vanno visti giacenti su un asse normale a detto piano e passante per  $\delta = 0$ ; siamo in uno spazio a tre dimensioni. Inoltre, per  $n_1$  e  $n_2$  crescenti (ad es. da 5 in poi), le potenze (come noto) crescono, e si ha un ventaglio di curve nel piano:  $\delta, 1 - \beta$ . Si arriva a una famiglia di curve che, nel caso più semplice di test bilaterale: se  $\pi_1 = .5$  sono simmetriche, a V; se  $\pi_1 < .5$  hanno andamento tendenziale ad L per la riduzione del ramo di destra; se  $\pi_1 > .5$  tendenziale a J per la riduzione del ramo di sinistra. Se  $\pi_1$  e  $\delta$  fossero variabili continue avremmo delle superfici. Sono elementi iniziali sulla potenza del test. (Nel caso in cui il lettore sia interessato all'argomento, si possono fornire gratuitamente i programmi per ottenere i risultati).

### 3.1 IL RUOLO DI $\pi$ NEL TEST

Tornando all'uso del test, trattiamo la scelta dello stato di Natura (sotto  $H_0$ )  $\pi$ , che deve essere fissato dal ricercatore, e che nel seguito del lavoro assumerà ancora un ruolo fondamentale.

Consideriamo due situazioni 'limite' di ricerca:  $\pi$  prossimo a 0 e  $\pi$  prossimo a 1.

Per la prima immaginiamo il caso in cui si voglia sperimentare l'efficacia di una strategia di persuasione in un consesso ostile all'argomento trattato, con probabilità  $\pi$  dei soggetti favorevoli prossima a 0. E' il caso di un ricercatore che, ad esempio, voglia convincere donne appartenenti ad una associazione femminista ad accogliere favorevolmente un'iniziativa di tipo maschilista. In tale ambiente, in cui argomenti di quel tipo sono sempre stati poco graditi, difficilmente le opinioni muteranno in maniera significativa. Vengono selezionati un gruppo di controllo  $n_1 = 18$  e sperimentale  $n_2 = 20$  che producono  $a_1 = 1$  e  $a_2 = 2$ ; l'ipotesi è monodirezionale sinistra: dopo il trattamento, per avere successo, la proporzione deve aumentare, e il test deve dare una  $dp_s = p_1 - p_2 < 0$ . In questa ottica, con i dati proposti risulta  $dps = -.044$ . Per quanto detto sopra, ponendo per la massa un  $\pi$  molto piccolo, .05, si determina la distribuzione campionaria della variabile  $dp$  sotto  $H_0: \pi_1 = \pi_2 = \pi = .05$  (da cui  $\gamma_1(dp) = .07$  e  $\gamma_2(dp) = .4, \sigma^2(dp) = .005$ ), e successivamente, sulla coda di sinistra, il p-value: frequenza cumulata fino a  $dp_s$ . La differenza è sì negativa, ma il p-value è .357, maggiore di  $\alpha = .05$ , quindi il trattamento non ha avuto successo.

Oppure, nello stesso ambiente, si consideri un'iniziativa di tipo femminista. Il trattamento, essendo in armonia con le opinioni del consesso, avrà facilmente successo in quanto gradito:  $\pi$  prossimo a 1. Con le stesse ampiezze campionarie  $n_1 = 18, n_2 = 20$ , e ad es.  $a_1 = 10$  (leggera propensione in assenza di trattamento) e  $a_2 = 19$  (notevole

consenso ai contenuti del trattamento), si ha  $dp_s = -.394$ . Ponendo  $\pi = .95$  ( $\gamma_1(dp) = -.07$ ,  $\gamma_2(dp) = .4$ ,  $\sigma^2(dp) = .005$ ), in termini di p-value si ottiene .0000008: il test è significativo, il trattamento ha avuto successo. Conformemente all'ipotesi e per evidenziare lo scopo dell'intervento, negli esempi la  $dp_s$  si è fatta risultare volutamente negativa.

E' chiaro il ruolo che assume  $\pi$ <sup>8)</sup>. Notiamo inoltre che i risultati dei due test **statistici** coincidono con ciò che è **intuitivamente** prevedibile, dati gli stati di Natura costituiti da opinioni così estremizzate: l'indisponibilità o la disponibilità ad accogliere i contenuti dei trattamenti per indurre al cambiamento di opinione. (Tale considerazione, più in generale, suggerisce cautele nell'uso dei risultati dei test in ambienti diversi da quelli di sperimentazione. In particolare, ad esempio, una stessa analisi della varianza, eseguita a Toronto e a Nairobi, darebbe gli stessi risultati?)

Il ricercatore può fare riferimento:

a) al valore dello stimatore (2) fornito dal campione, soprattutto per confronti con i risultati forniti dall'approssimazione mediante la Normale.

La (2) viene comunemente usata nella (1) per standardizzare la variabile  $p_1 - p_2$  con riferimento alla Normale; nel caso attuale si utilizza per costruire la distribuzione campionaria partendo dalla (3).

b) ad una valutazione basata sull'intervallo di confidenza di  $\pi$  (da programma) ottenuto tramite lo stimatore (2) e, in quell'ambito, ponendo attenzione ad es. al valore del parametro che rende più (o meno) prudentiale il test.

Altri Autori [1], propongono di respingere  $H_0$  se, per tutti i valori (discreti) di  $\pi$  contenuti nell'intervallo, gli esiti del test sono concordi nel respingerla.

c) al concetto di significatività (o non) assoluta e condizionata.

Vedremo l'estensione del punto b) al caso di test assolutamente significativo (o non), e al caso di condizionamento, che fornirà alcuni spunti di riflessione.

d) soprattutto a sue esigenze o convinzioni scientifiche.

In tal caso l'operatore sottopone a test l'ipotesi  $H_0$  ( $\pi_1 = \pi_2 = \pi$ ), fissando uno stato di Natura  $\pi$  da lui proposto.

e) a informazioni precedentemente acquisite.

I punti a) e b) non hanno bisogno di approfondimenti e commenti. Nel lavoro interessa soprattutto il punto c) che sarà visto in seguito. Per i casi d) ed e) seguono due esempi.

Un ricercatore vuole provare l'efficacia di una sostanza stimolatrice della memoria, da lui preparata, in un insieme di cavie. In base a **nozioni teoriche** in suo possesso e suoi obiettivi da raggiungere ritiene di far riferimento ad un sistema di ipotesi,  $H_0: \pi_1 = \pi_2 = .8$  ed  $H_1: \pi_1 > \pi_2$ . Ipotesi monodirezionale destra, per  $\alpha = .025$ , Tabella 1.

Dispone di due gruppi limitati di soggetti,  $n_1 = 15$  ed  $n_2 = 10$  (numerosità della Tabella scambiate), sperimentale e di controllo. Il numero dei successi, soggetti che reagiscono positivamente alla sostanza, è  $a_1 = 14$  e  $a_2 = 5$ , danno luogo alle proporzioni  $p_1 = 14/15 = .93$  e  $p_2 = 5/10 = .5$  con una  $dp_s = .43$ .  $\pi = .8$  non è in Tabella, ma possiamo riferirci a  $\pi = 1 - .8 = .2$  (altro scambio in Tabella, tra  $\pi$  e  $1 - \pi$ ) e i valori critici sono letti direttamente. La  $dp_s = .43 > dp'' = .367$ , il test è significativo per  $\pi = .8$ .

Altro esempio: valore noto per **informazioni precedentemente acquisite**.

Un gruppo politico in base a sondaggi precedenti lontani nel tempo, notizie ufficiose o valutazioni indirette su caratteri correlati, ritiene di essere presente in una determinata località nella percentuale di circa il 6%, il che equivale a dire in una proporzione  $\pi = .06$ . Questo è il valore d'interesse, lo stato di natura di riferimento per il test; a questo punto si procede come di consueto.

### 3.2 IL RUOLO DEI SOTTOINSIEMI DI $\pi$

Per semplicità consideriamo i valori di  $\pi$  equiprobabili.

Anche direttamente dalle Tabelle 1 e 2 si era notato che per le varie coppie di numerosità campionarie, i valori critici superiori  $dp''$  variano da un valore inferiore ad un valore superiore; in Tabella 1 per  $n_1 = 10$  e  $n_2 = 15$   $dp''$  varia da .233 a .433, mentre i valori di  $dp'$  da un valore superiore ad un valore inferiore, da -.167 a -.433. Quindi, dato  $\alpha$ , nel caso di test monodirezionale destro, se una differenza sperimentale (positiva)  $\geq dp''$  più elevato, il test è sempre significativo, per ogni  $\pi$ ; se è  $< dp''$  meno elevato, il test sempre non significativo, per ogni  $\pi$  (a parte il dettaglio grossolano dei valori). E analogamente per le code a sinistra, se una differenza sperimentale (negativa)  $\leq dp'$  meno elevato, il test è sempre significativo; se è  $> dp'$  più elevato, il test è sempre non significativo.

Per facilitare la comprensione degli aspetti che vogliono essere messi in evidenza nel presente lavoro, con l'aiuto delle Tabelle facciamo un esempio introduttivo all'aspetto che sarà sviluppato in seguito per qualsiasi valore di  $\pi$  e delle numerosità  $n_1$  ed  $n_2$ . Tabella 1,  $\alpha = .025$  su una coda. Dati  $n_1 = 15$ ,  $n_2 = 25$ , se l'ipotesi è monodirezionale

destra il valore critico per  $\pi = .05$  è .160, per  $\pi = .5$  è .333 per  $\pi = .95$  è .133. Per valori  $a_1 = 8$  e  $a_2 = 4$ , il valore sperimentale  $dp_s = .373$  è **sempre significativo**, perché per tutti i valori di  $\pi$  è superiore ai valori critici della coda destra. Invece per  $a_1=6$  e  $a_2=4$ ,  $dp_s = .240$  è **significativo**, in quanto superiore ai valori critici della coda destra, **solamente** nei sottoinsiemi per valori di  $\pi = .05, .1, e .85$  (.240>.227), **.9** (.240>.187), **.95** (.240>.133).

In un caso più generale, usando il DPT che ci permette di considerare numerosità  $n_1$  ed  $n_2$  non previste in Tabella, dato un certo risultato sperimentale ed il livello di significatività  $\alpha$ , il ricercatore può vedere in maniera più dettagliata per quali valori di  $\pi$  si ottengono risultati significativi data una  $dp_s$  mediante il confronto con i valori critici, come segue.

Esempio. Per  $n_1 = 12, n_2 = 18, a_1 = 9, a_2 = 5$ , si ha  $dp_s = .4722$ . Nel caso di estrazione con reimmissione, per  $\alpha = .025$  ed ipotesi monodirezionale destra, il test è sempre significativo.

$\pi$	$dp_s$			
.05	.1944			
.1,	.2777			PARTE DESTRA
.15,	.3055			
.2	.3333			
.25	.3611			
.3	.3611			
.35	.3888			
	”			
.....				
.6	.3888			
.65	.3611			
.7	.3611			
			$\pi$	$dp_s$
			.72	.3611
			.722	.3611
			.723	.3333
.75	.3333			
.8	.3055			
.85	.2777			
.9	.2500			
.95	.1944			

Il dettaglio<sup>4)</sup> dei valori di  $\pi$  può essere migliorato come risulta nella PARTE DESTRA del prospetto (.72, .722, .723) e si può arrivare a valori limite a discrezione dell'operatore. A sinistra si nota che il test è **significativo per tutti i valori di  $\pi$** ; ma (come nell'esempio precedente) se il numero dei successi fosse  $a_1 = 10$  e  $a_2 = 9$ , si avrebbe  $dp_s = .3333$  significativo solo per  $\pi = .05, .1, .2$ , oppure per  $\pi = .75, .8, .85, .9, .95$ .

Il procedimento seguito, in sostanza fornisce un prospetto di risultati: i valori critici, al variare dei valori di  $\pi$ . Forse è utile corredare un risultato con tale caratterizzazione, non prevista quando si usa la Normale o in altri test. Notiamo che  $\pi$ , da causa di perplessità (parametro di disturbo), assume un ruolo di elemento informativo sul test.

### 3.3 SIGNIFICATIVITÀ ASSOLUTA E CONDIZIONATA

Pertanto dato  $\alpha$  e le numerosità  $n_1$  ed  $n_2$ , al variare di  $\pi$ , l'escursione dei valori critici permette di fare la seguente considerazione.

Un valore sperimentale che risulti significativo per ogni  $\pi$ , potrebbe essere considerato “assolutamente” (o incondizionatamente) significativo, mentre un valore che risulti significativo solo per un sottoinsieme di valori di  $\pi$  “relativamente” (o condizionatamente) significativo.

Definizione. Per un dato  $\alpha$ , un valore sperimentale  $dp_s$  che risulti significativo (o non) per ogni  $\pi$ , è “assolutamente” (o incondizionatamente) significativo (o non), mentre un valore che risulti significativo solo per un sottoinsieme di valori di  $\pi$ , è “relativamente” (o condizionatamente) significativo.

Gli esempi che seguono faranno riferimento al concetto di significatività assoluta o condizionata; ma la trattazione risulta più chiara utilizzando il p-value. Nei prospetti che seguono, i valori di  $\pi$  variano con passo .05, ma si può spingere l'analisi dettagliando di più tali valori, come già detto.

Consideriamo ora un esperimento che produca una differenza sperimentale molto grande in valore assoluto, negativa e tale che, dato  $\alpha$ , cada sempre in zona di rifiuto al variare di  $\pi$ , caso di significatività assoluta.

Due gruppi:  $n_1 = 15$  e  $n_2 = 20$ , se  $a_1 = 5$  e  $a_2 = 13$  risulta  $dp_s = -.317$  che, al variare di  $\pi = .05, \dots, .95$  (è bene ricordare che per ogni valore di  $\pi$  è sempre  $H_0: \pi_1 = \pi_2 = \pi$ ) in termini di p-value dà:

p-value	.000	.00	.004	.001	.017	.023	.028	.032	.034	.036	.035	.033	.03	.026	.02	.013	.006	.002	.00
$\pi$	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5	.55	.6	.65	.7	.75	.8	.85	.9	.95

Posto  $\alpha = .05$  il p-value risulta sempre inferiore al livello di significatività. In sostanza con quelle numerosità e risultati sperimentali, per quel valore di  $\alpha$  il test è assolutamente significativo (aspetto già visto, ma qui si fa riferimento al p-value).

Oppure un esperimento che dia luogo ad una  $dp_s$  prossima a 0 o nulla, tale da non cadere mai in zona di rifiuto al variare di  $\pi$ . Con le ampiezze precedenti se  $a_1 = 9$  e  $a_2 = 11$ ,  $dp_s = .05$ , dà:

p-value	.266	.311	.355	.381	.387	.39	.397	.4	.4	.4	.4	.4	.4	.396	.396	.389	.362	.333	.296
$\pi$	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5	.55	.6	.65	.7	.75	.8	.85	.9	.95

Posto  $\alpha = .025$ , il p-value è sempre superiore ad  $\alpha$ : il test non è mai significativo.

Nei casi esemplificati, dato  $\alpha$ , il ricercatore che non ha informazioni su  $\pi$ , ma dispone dei prospetti, rifiuta o non rifiuta l'ipotesi nulla, anche senza fissare tale valore.

**In generale, nell'incertezza circa lo stato di Natura  $\pi$ , nel caso di assoluta significatività (o non) si potrà dar credito, e con il rischio stabilito  $\alpha$ , al successo (o non) del test.**

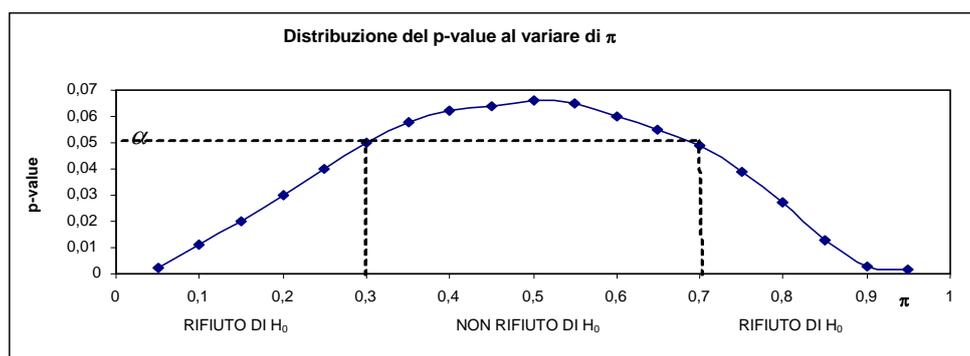
Ma nella maggior parte dei casi avviene che, per un sottoinsieme di valori di  $\pi$ , si abbiano risultati significativi mentre per il complementare si abbiano risultati non significativi.

Consideriamo la prova di un farmaco utilizzando due gruppi indipendenti, sperimentale  $n_1 = 15$  e di controllo  $n_2 = 20$ , con numero di successi (guarigioni) rispettivamente  $a_1 = 10$  e  $a_2 = 8$ . La differenza sperimentale è  $dp_s = .27$ . Il prospetto che segue riporta il p-value al variare di  $\pi$ , con  $\alpha = .05$ .

p-value	.0022	.011	.02	.03	.04	<b>.05</b>	.058	.062	.064	.066	.065	.06	.055	<b>.049</b>	.039	.027	.013	.003	.0016
$\pi$	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5	.55	.6	.65	.7	.75	.8	.85	.9	.95

I valori di  $\pi$  in neretto portano al rifiuto di  $H_0$ .  $\pi = .3$  e  $\pi = .7$  delimitano i sottoinsiemi di rifiuto.

Graficamente, al variare di  $\pi$  si ha [17]:



Per ampiezze  $\leq 30$ , il problema si risolve per valori discreti di  $\pi$ , utilizzando il programma come già illustrato. Il procedimento è laborioso ed approssimativo, ma potrebbe essere programmato per avere una maggior rapidità e dettaglio dei valori di  $\pi$ ; in sostanza, dati  $n_1, n_2, a_1$  e  $a_2$ , ottenere immediatamente la distribuzione del p-value, quindi dato  $\alpha$ , trovare gli estremi dei sottoinsiemi.

Per  $n > 30$  (Normale), si può trovare un procedimento tale da fornire mediante formule gli estremi dei sottoinsiemi. Per una interessante soluzione (in R) contattare: [michele.grassi@phd.units.it](mailto:michele.grassi@phd.units.it).

Un primo commento a tale distribuzione è il seguente. Poniamoci nel caso in cui il ricercatore **non abbia informazioni** per fissare  $\pi$ , e lo scelga a caso. Avrà una probabilità di rifiutare l'ipotesi  $H_0$ , approssimativamente data dalla somma delle ampiezze degli intervalli di  $\pi$ , compresi da valori prossimi a 0 fino a .3, e da .7 fino a valori prossimi a 1, cioè circa  $.3 + .3 = .6$ . In sostanza la probabilità che il trattamento dia luogo all'effetto sperato è valutata dalla somma delle ampiezze degli intervalli dei valori di  $\pi$ , come in figura. E' bene chiarire però che questa ampiezza non dà la probabilità che  $\pi$  ivi scelto a caso sia quello giusto: è solo la probabilità che scegliendo a caso un  $\pi$ , questa porti al rifiuto di  $H_0$ ; quindi, nel caso di incertezza circa il valore di  $\pi$ , il ricercatore, che di solito è interessato al rifiuto di  $H_0$ , non può affidarsi alla probabilità dei valori di  $\pi$  che portano al rifiuto. Invece, il ricercatore che **ha informazioni** seppur vaghe sul valore di  $\pi$ , disponendo del prospetto o del grafico, ha la possibilità di osservare e controllare se il valore da lui fissato sia sostenuto da un sottoinsieme più o meno ampio che lo comprenda; cioè se la valutazione di  $\pi$  da lui decisa rientra, meglio ancora se è centrata, in un sottoinsieme in cui il test sia o meno significativo. Maggiore è l'ampiezza del sottoinsieme intorno allo stato di Natura  $\pi$  da lui scelto, più il ricercatore è rassicurato circa l'aver rifiutato  $H_0$ , perché in quell'intorno tutti i valori di  $\pi$  portano al rifiuto; o viceversa se ne è fuori. Nell'esempio, siamo nel primo caso se il ricercatore ha informazioni per ritenere lo stato di natura  $\pi$  pari a circa .8, .85; siamo nel secondo caso se pari a circa .45, .5.

L'osservazione del grafico offre lo spunto per alcune riflessioni.

Un elemento informativo è rappresentato dal sottoinsieme in cui è compreso il valore fissato di  $\pi$  che porta al rifiuto o al non rifiuto: alternative sostenute dall'ampiezza dei sottoinsiemi in cui è compreso  $\pi$  stesso. Il ricercatore, che ha sempre come obiettivo l'efficacia del trattamento, potrebbe non essere tanto interessato al valore di  $\pi$  ma solo all'intervallo (al o ai sottoinsiemi) dei valori di  $\pi$  che portano al rifiuto. E da tale procedura potrebbe trarre la conclusione che se  $\pi$  assume valori compresi entro certi limiti il risultato sperimentale porta al rifiuto di  $H_0$ ; o viceversa per il non rifiuto. Il ricercatore potrà e dovrà riflettere sulla teoria e sul modello proposti, soprattutto nel caso in cui il test risulti significativo in un sottoinsieme di valori di  $\pi$ , non previsti, ma ritenuti sensati.

Con tale approccio si dà al risultato dell'esperimento la fiducia di base che merita, nel senso che è l'elemento concreto, certo, di cui dispone il ricercatore, mentre si lascia a questo la possibilità di una riflessione sul sottoinsieme di significatività, che può anche orientarlo nella scelta del valore  $\pi$ , stato di Natura, (nella formulazione di una teoria, e/o del modello), della cui scelta rimane protagonista.

Pertanto con le ampiezze  $n_1, n_2$  ed i numeri dei successi  $a_1$  e  $a_2$ , si costruisce il prospetto o il grafico; disporre di tale distribuzione significa avere il quadro completo delle possibilità di interpretazione, dato lo stato di fatto fornito dall'esperimento, che è l'unico riferimento certo. Successivamente viene considerato  $\pi$ . L'elemento informativo, il supporto per una riflessione su  $\pi$ , è rappresentato dal prospetto del p-value. Il ricercatore avendo sotto gli occhi il quadro della situazione, deciderà in base alla teoria e ai suoi obiettivi. Essendo interessato al rifiuto dell'ipotesi nulla, potrà rivolgere la sua attenzione al sottoinsieme dei valori di  $\pi$  che danno luogo ai p-values di valore inferiore ad un limite  $\alpha$  da lui prefissato.

L'approccio mediante il prospetto del p-value al variare del parametro potrebbe essere utilizzato anche nel caso di altri test considerando una gamma di valori possibili e sensati del parametro stesso, e non equiprobabili.

Riassumendo, con i dati del campione possiamo considerare i seguenti casi. Significatività assoluta: il test è significativo per ogni  $\pi$ ; assoluta non significatività: il test non è significativo per ogni  $\pi$ . In questi due casi il ricercatore può decidere senza perplessità.

Oppure il test è significativo per sottoinsiemi di valori di  $\pi$ , è questo il caso di interesse. Il ricercatore avrà modo di vedere se i sottoinsiemi siano coerenti con le proprie idee sulla teoria di base, o comunque se i valori compresi siano accettabili per il tipo di ricerca.

Osservando i sottoinsiemi dei valori del parametro per i quali l'esperimento risulta efficace, nasce l'idea che forse è opportuno abbandonare la pretesa di voler valutare il parametro per realizzare il test (provare un'ipotesi predefinita), e invece accettare ciò che l'esperimento offre in termini di significatività tramite detti sottoinsiemi e da questi trarre conclusioni. Questo viene offerto dalla distribuzione o dal prospetto del p-value, e ci permette la seguente **provocazione**: ribaltare il discorso relativo al test. Anziché proporre il parametro sotto ipotesi e poi vedere il risultato del test *in termini di significatività* e quindi 'rifiutare' o non l'ipotesi nulla, dare fiducia alle osservazioni e prendere atto del risultato del test *in termini di sottoinsiemi* che portano al rifiuto al variare di  $\pi$ .

**Conclusione.** Con tale approccio è dopo l'esperimento che maturano e forse nascono idee sullo stato di Natura, che comunque rimarrà sempre ignoto; le posizioni sembrano ribaltate, non più parametro-risultato del test, ma risultato del test-parametro, e con esse la logica del procedimento. Questa opportunità è offerta dal prospetto o dalla distribuzione sopra visti: un esperimento è risultato significativo per un sottoinsieme di valori del parametro, valori insospettati, ottenuti un po' per 'caso', vale la pena di riflettere sul loro significato. Se è vero che con un esperimento si cerca la conferma di una teoria, è anche vero che l'esperimento può suggerirne una nuova.

#### 4.1 ESTRAZIONE SENZA REIMMISSIONE, CAMPIONI INDIPENDENTI

Il caso di estrazione senza reimmissione è il più usato nel campionamento da popolazioni finite, soprattutto nel caso di masse limitate e piccoli campioni, e genera universi più rappresentativi.

Data una popolazione di numerosità  $N$  divisa in due modalità A e B con frequenze assolute  $N\pi$  ed  $N - N\pi$ , interi, presentiamo il caso di campioni indipendenti molto utile per la didattica perché tramite computer è possibile seguire la tendenza al caso di estrazione con reimmissione al diminuire della frazione sondata f. s. =  $n/N$  ( $N$  crescente).

Per due campioni indipendenti, di ampiezza  $n_1$  ed  $n_2$  e numero di successi  $a_1$  e  $a_2$ , applicando la distribuzione ipergeometrica abbiamo:

$$fr(p_1)fr(p_2) = \frac{\binom{N\pi}{a_1} \binom{N(1-\pi)}{n_1 - a_1} \binom{N\pi}{a_2} \binom{N(1-\pi)}{n_2 - a_2}}{\binom{N}{n_1} \binom{N}{n_2}}. \quad (10)$$

Tale approccio può essere usato in casi particolari tenuto conto delle seguenti considerazioni.

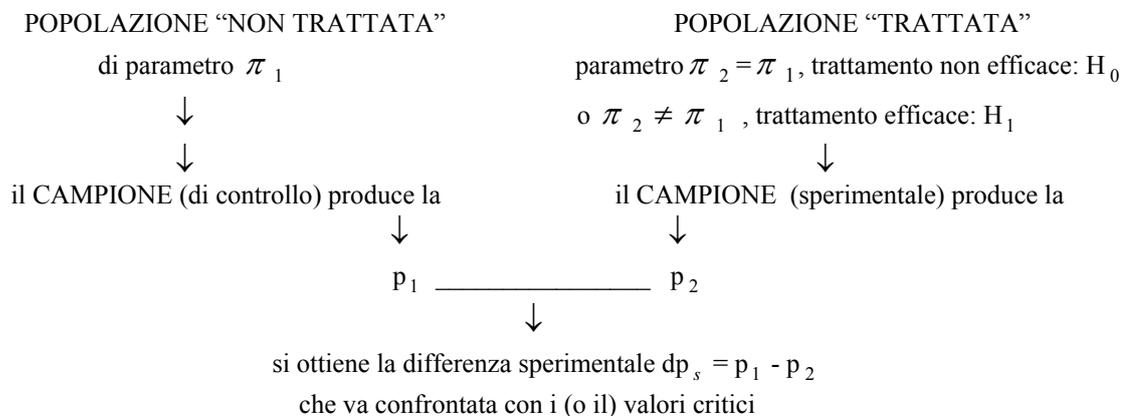
In alcuni tipi di indagine per motivi di riservatezza o per limitare la pubblicità di una iniziativa di ricerca è opportuno far riferimento a masse limitate da cui estrarre campioni.

Siamo nel caso di esperimenti in cui si voglia verificare ad esempio l'efficacia di una campagna elettorale o pubblicitaria, fondata su una strategia di persuasione innovativa e prima di lanciarla nel campo nazionale a cui è destinata. Inoltre, in alcuni settori di ricerca, un trattamento effettuato su un campione, date le circostanze in cui si svolge e la cura con cui è possibile procedere nel caso di un gruppo limitato, non fornisce le stesse informazioni date da un campione estratto da una popolazione sulla quale preventivamente sia stato applicato un trattamento, che di solito si realizza in maniera meno accurata, più dispersiva.

Ad esempio, un professore di una Facoltà ha una classe di 300 studenti che danno risultati mediocri e vuole sperimentare un nuovo metodo di insegnamento. Le lezioni si svolgono naturalmente dalla cattedra; per esperienza ha constatato che se illustra gli stessi argomenti a un gruppo limitato, in privato, sembra che tutti capiscano; i due interventi sono sullo stesso piano? O un diverso livello di empatia e componenti ambientali giocano un ruolo non trascurabile. Obiettivamente il 'trattamento' dà risultati diversi nei due casi: operando su un gruppo limitato si ha un effetto qualitativamente (e quantitativamente) migliore, rispetto a quello che si ottiene rivolgendosi alla classe. Il professore per provare l'efficacia del nuovo metodo d'insegnamento deve applicare il trattamento all'intera classe anziché ad un campione di studenti [13]. E in tutte le indagini che prevedono un coinvolgimento attivo dei soggetti interessati, operando su un campione si ha un effetto diverso da quello che si ottiene "trattando" direttamente la popolazione che è il vero obiettivo (starà poi al campione con le regole offerte dalla metodologia rappresentarla).

Ma anche nella ricerca biologica, considerando il mondo animale nel caso di specie rare o viventi in particolari zone e magari in via di estinzione, per indagare sul loro comportamento allo stato libero o per collaudare un intervento di salvaguardia, il trattamento viene fatto nel loro ambiente e non su un gruppo (campione) di cavie allevate dall'uomo o comunque in cattività.

Tali considerazioni prevedono la seguente procedura. Da una massa limitata o popolazione, si estrae un campione di controllo che produce la  $p_1$ , poi si ricostituisce la popolazione rimettendovi il campione, successivamente si interviene con il trattamento sulla popolazione quindi, considerando immutate altre componenti, si estrae il secondo campione che produce la  $p_2$ . In un prospetto si ha:



I due campioni sono indipendenti (la popolazione numericamente non cambia), quindi la (10).

L'obiezione, secondo cui il ricostituire la popolazione rimettendo i soggetti del campione di controllo nella stessa, altererebbe la sua struttura perché essi conservano il ricordo del trattamento, è sì vera, ma forse il toglierli l'avrebbe

alterata ancora di più nel caso di popolazioni di numerosità limitata; per questa ragione è preferibile tenere costante la numerosità  $N$ .

I valori critici sono determinati, sotto  $H_0$ , mediante la distribuzione casuale delle  $dp$  e dato il parametro  $\pi$ .

La procedura dovrebbe essere applicata in tutte le indagini in cui si ritenga siano presenti aspetti o componenti che possano differenziare i due tipi di intervento: sul campione o sulla popolazione.

Esempio: una ditta fornisce un prodotto di utilizzazione familiare; date le scarse vendite, vuole sperimentare un nuovo tipo di slogans e cartelli per lanciarsi poi in una campagna pubblicitaria. La prova viene svolta con le seguenti modalità. L'incaricato dell'indagine sceglie opportunamente una località, un centro abitato, ad esempio di  $N = 200$  famiglie e posto come riferimento  $\pi = .12$ , noto da statistiche precedenti, decide per l'estrazione senza reimmissione e f.s. = .1. Da liste Comunali sorteggia un campione di 20 famiglie (non disponibili per un esperimento Prima - Dopo) e ottiene una  $p_1 = .15$  (3/20) di preferenza per quel prodotto. Rimesso il campione nella massa, applica su quel centro abitato il trattamento e successivamente estrae un nuovo campione ad esempio di 20 famiglie, che produce una  $p_2 = .45$  (9/20). La differenza sperimentale  $dp_s = -.3$ . Dato  $\alpha = .05$  ed ipotesi alternativa bilaterale (è incerto sull'esito del trattamento) i risultati forniti dal programma sono: p-value = .0023, che già basterebbe perché  $.0023 < \alpha/2 = .025$ ; inoltre i valori critici, inferiore =  $-.25$  e superiore =  $.25$ , confermano che essendo  $-.3 < -.25$ , il test è significativo. Dato quel  $\pi$  e quel livello  $\alpha$  quel tipo di pubblicità sarebbe efficace.

Il procedimento appena visto offre lo spunto per le seguenti considerazioni di validità più generale.

E' la stessa cosa "trattare" una popolazione e poi estrarre un campione, oppure estrarre un campione e poi applicare al gruppo estratto un trattamento? Lo scambio è irrilevante circa la somministrazione del trattamento e dei suoi effetti? Nel caso di trattamento sul gruppo i risultati del test possono portare ad esempio a  $\pi_1 \neq \pi_2$ ; ma il gruppo

trattato può ritenersi un campione casuale estratto da una massa caratterizzata da  $\pi_2$ ? Inoltre, dopo l'esperimento sul gruppo, circa il trattamento da applicare alla popolazione, non si ha qualche dubbio (come e quanto)?

Si può aggiungere che se si estrae un campione casuale dopo un trattamento sulla popolazione, si può parlare di 'campione' a tutti gli effetti, altrimenti si dovrebbe parlare di un 'gruppo' selezionato e trattato, e con le perplessità conseguenti.

Sotto  $H_0$  la distribuzione della  $dp$  ha:  $E(dp) = 0$  e  $var(dp) = \pi(1 - \pi)(N/n_1 + N/n_2 - 2)/(N - 1)^9$ .

La varianza è minore del caso con reimmissione che utilizza la (3), pertanto computazionalmente i valori critici e nella maggioranza dei casi risultano più vicini tra loro rispetto al detto caso; e ancor di più al crescere della frazione sondata; viceversa per f.s. .0001 si ha sostanzialmente coincidenza con il caso con reimmissione. La (10) privilegia l'ipotesi  $H_1: \pi_1 \neq \pi_2$ , il test è meno conservativo.

-In alcune occasioni, e per certi tipi di indagine, potrebbe verificarsi che il risultato di un trattamento su un campione sia tale da non rientrare addirittura nella distribuzione campionaria dei risultati, ottenuta mediante l'estrazione casuale da una popolazione a cui sia stato applicato il trattamento: è questo il pericolo pratico che si corre quando ci si allontana dalla procedura illustrata, oltre alle perplessità di fondo circa lo scambio tra gli oggetti del trattamento, campione e popolazione. Tale procedura presenta l'inconveniente di dover essere applicata a popolazioni limitate e, per avere rappresentatività più estesa, dovrebbe trovare completamento considerando più siti di indagine opportunamente scelti. Con la **meta-analisi**, il cui obiettivo è quello di integrare una serie di risultati ottenuti da singoli studi, si potrebbe raggiungere tale risultato.

## 4.2 GRUPPI NON INDIPENDENTI

Seguendo l'approccio consueto di due gruppi e due trattamenti, caso di confronti ad esempio tra medicinali, regimi alimentari, ecc. (qualora non intervengano le componenti ambientali o comportamentali del tipo visto), per piccoli gruppi  $n_1$  ed  $n_2$  estratti da una massa limitata, si ha:

$$\begin{aligned} \text{fr}(p_1)\text{fr}(p_2 | p_1) &= \binom{N\pi}{a_1} \binom{N(1-\pi)}{n_1 - a_1} \binom{N\pi - a_1}{a_2} \binom{N(1-\pi) - n_1 + a_1}{n_2 - a_2} \bigg/ \binom{N}{n_1} \binom{N - n_1}{n_2} = \\ &= \binom{n_1}{a_1} \binom{n_2}{a_2} \binom{N - n_1 - n_2}{N\pi - a_1 - a_2} \bigg/ \binom{N}{N\pi}. \end{aligned} \quad (11)$$

Dalla prima delle (11) a destra dell'uguale e al denominatore, si nota che dopo l'estrazione del primo gruppo la numerosità della massa cambia da  $N$  a  $N - n_1$ . Dalla seconda si evidenzia che scambiando gli indici il risultato non cambia. Compatibilmente con la positività dei coefficienti, le (11) sono valide da 1 a più gruppi.

La distribuzione ha  $E(p_1 - p_2) = 0$  e  $\text{var}(p_1 - p_2) = N \pi (1 - \pi)(1/n_1 + 1/n_2)/(N-1)$ , maggiore della precedente perché la covarianza tra  $p_1 = a_1/n_1$  e  $p_2 = a_2/n_2$  è negativa<sup>10</sup>.

In assenza di trattamenti tutte le differenze sono casuali e siamo sotto l'ipotesi  $H_0$ . Con riferimento ad  $H_0$  vengono determinati i valori critici, che risultano più distanti tra loro (sostanzialmente coincidenti con quelli dell'estrazione con reimmissione). Viene privilegiata l'ipotesi  $H_0: \pi_1 = \pi_2$  il test è più conservativo rispetto al caso precedente.

Estratti i gruppi ed effettuati i trattamenti questi possono dar luogo ad una differenza tale che se cade in zona di rifiuto viene attribuita (con rischio  $\alpha$ ) ad una componente sistematica.

Notiamo che una differenza ottenuta tra "gruppi" trattati viene discriminata in base a valori critici ricavati da una distribuzione di differenze ottenuta mediante "campioni" casuali.

Il ricercatore valuterà la presenza delle circostanze per adottare il primo o il secondo procedimento.

-Osservazione. Poniamo il caso più semplice di campione di controllo (non trattato) e gruppo sperimentale (trattato).

Il test è strutturato su  $H_0 (\pi_1 = \pi_2 = \pi)$  e sotto tale ipotesi viene determinata la distribuzione della variabile  $dp = p_1 - p_2$  da cui si ricavano i valori critici; tale distribuzione è ottenuta da campioni **casuali** (estratti dalla stessa popolazione non trattata), in conseguenza si ha rappresentatività sia per  $p_1, p_2$  e  $dp$ , che per i valori critici  $dp'$  e  $dp''$ .

Ed anche la differenza sperimentale  $dp_s = p_1 - p_2$  viene considerata come determinazione di un valore casuale appartenente alla distribuzione sotto  $H_0$  (sia che valga sia che non valga  $H_0$ ), salvo poi a vedere se è possibile rifiutare tale ipotesi. Ma mentre nel caso di non significatività per la  $dp_s$  si ritiene che sia il campione di controllo sia il campione sperimentale provengano dalla stessa popolazione (sotto  $H_0$  il parametro non ha subito variazioni) e quindi non dovrebbero esserci perplessità; invece nel caso di significatività - che dipende da  $\alpha$  - occorre tener presente che la  $dp_s$  sarebbe ottenuta mediante la componente  $p_2$  prodotta da un gruppo non estratto casualmente da una popolazione trattata e caratterizzata da una  $\pi_2 \neq \pi_1$ : il gruppo sperimentale non è un campione casuale di tale popolazione: manca di rappresentatività. In altre parole la  $dp_s$  risulta dalla differenza tra una determinazione **casuale** rappresentativa ( $p_1$ ) e **non casuale** non rappresentativa ( $p_2$ ). E per di più, è confrontata con un valore critico proveniente da campioni casuali (ambidue rappresentativi) sotto  $H_0$ . Quindi nel caso di significatività, sia la determinazione della  $dp_s$  sia il confronto con il valore critico fanno nascere qualche perplessità. L'osservazione si potrebbe fare anche nel caso più generale del piano degli esperimenti; in cui i valori critici vengono determinati mediante la distribuzione campionaria (casuale) sotto  $H_0$ . Generalmente questi ultimi aspetti non vengono segnalati.

## 5.1 ASPETTI STATISTICI CONSIDERATI PER LA COSTRUZIONE DEL PROGRAMMA DPT

Il programma Dpt.exe, Fortran, può essere fornito gratuitamente. E' stato costruito come descritto nella Sezione 2.1- e riporta i risultati ottenuti dalla (3) e dalla (10), mentre per quelli ottenibili dalla (11) pur esistendo un programma specifico, come già detto si può far riferimento ai risultati dell'estrazione con reimmissione.

**Nel programma il parametro  $\pi$  è indicato con P, quindi qui viene indicato con P.**

Nel campionamento casuale semplice da popolazione finita di numerosità  $N$ , caso di estrazione senza reimmissione, gioca un ruolo importante la frazione sondata f.s. =  $n/N$  ([8] Cochran pag. 21, [18] Sarndal pag. 66), coincidente peraltro con la probabilità di inclusione del primo ordine delle unità della popolazione, ([8] Cochran pag. 259, [18] Sarndal pag. 30,31). Nel lavoro [14] ( $n_1 = n_2 = n$ ), per porre i vari casi nella stessa situazione probabilistica, la frazione sondata era stata considerata fissa al variare di  $n$  ([6] Cicchitelli ed Altri pag. 57, 58), e considerata successivamente pari a .1, .01, .001, .0001. Aggiungiamo che per f.s. = .0001 si ha sempre la coincidenza computazionale in termini di valori critici con l'estrazione con reimmissione. Nel caso attuale si è seguita la stessa regola. Ma nel test che stiamo trattando, abbiamo due campioni estratti dalla stessa popolazione, e se i campioni hanno numerosità diverse, si hanno due frazioni sondate  $n_1/N$  e  $n_2/N$ . Per assicurare che dette frazioni almeno non superino un valore stabilito per ambedue i campioni, si è deciso di sceglierne una che limiti anche l'altra, è stata scelta  $n_2/N$ . Quindi nel caso di estrazione senza reimmissione viene richiesto di **non** operare con  $n_1 > n_2$ ; significherebbe inoltre rischiare una numerosità  $n_1$  troppo elevata rispetto ad  $N$ . Invece nell'estrazione con reimmissione non si hanno riferimenti limitativi e quindi è possibile fissare  $n_1 > n_2$ . Per quanto riguarda l'intervallo di confidenza, si tratta ancora di un procedimento nel discontinuo. E' alternativo a quello valido nel caso di in un

solo campione illustrato nell'articolo di Clopper & Pearson [7] e in Cramer [9], pag. 511 – 515. Fissato il livello di confidenza CONFL ( $= 1 - \alpha$ ) per ogni P (valori discreti, ascisse), il programma determina  $p'(P)$  e  $p''(P)$  (ordinate), estremi dei sottoinsiemi di dimensione  $\geq 1 - \alpha$  che seguono due andamenti per punti, uno superiore ed uno inferiore. I sottoinsiemi considerati nella procedura sono quelli che presentano frequenze sulle code  $\leq \alpha/2$ , gli altri vengono esclusi. Utilizzando la stima (2), le ascisse delle intersezioni tra la retta orizzontale della stima e i due andamenti, forniscono gli estremi dell'intervallo P1 e P2; nel caso di non coincidenze con i valori di P (discreti) sopra visti, il programma esegue una interpolazione tra i valori contigui di P interessati. La procedura è applicabile per:  $3 \leq n_1$  ed  $n_2 \leq 30$ , dato il CONFL. Ma per numerosità molto piccole potrebbe essere necessario diminuire il CONFL per accostare i due andamenti superiore ed inferiore e permettere la loro intersezione con la retta orizzontale della stima.

Può essere fornito il programma che utilizza la ipergeometrica (11); ed anche il programma per determinare solo l'intervallo di confidenza della P. Il programma stampa le coordinate dei punti del grafico, quindi l'operatore può determinare l'intervallo, dato un valore dello stimatore (2). Può anche essere fornito il programma DRT .exe relativo a matched groups, per numerosità  $n \leq 30$ , utilizzato nei lavori [12] e [15], corredato da istruzioni.

## 5.2 USO DEL PROGRAMMA DPT

### Per comprendere meglio quanto segue si consiglia di operare con il programma DPT.exe.

Le successive istruzioni di input prevedono inizialmente la scelta tra l'esecuzione di un test o la costruzione di Tavole di valori critici, dando: 0 o 1. Se si dà 0, successivamente occorre dare:

1)  $n_1$  e  $n_2$ : le numerosità dei gruppi; nel caso di estrazione senza reimmissione si consiglia di lavorare con  $n_1 \leq n_2$ .

2)  $a_1$  e  $a_2$ : numero dei successi nel primo e nel secondo gruppo,  $a_1 \leq n_1$ ,  $a_2 \leq n_2$ .

Il programma determina la differenza sperimentale dps e la stima (2) della P,  $stp = (a_1 + a_2) / (n_1 + n_2)$ .

In una fase iniziale, per rendere più semplice l'approccio il programma lavora mediante estrazione con reimmissione e utilizzando la stp, stima del parametro P. Quindi richiede il valore di:

$\alpha$ , area o probabilità ( $\leq .5$ ) di rifiutare  $H_0$  quando è vera; per un test bilaterale,  $\alpha$  viene diviso per due e  $\alpha/2$  viene riferito a ciascuna coda; si possono dare più valori di  $\alpha$  successivamente;

se si dà  $\alpha = 0$ , viene utilizzata come deviatore, e il programma dà l'area dal valore sperimentale dps in poi verso la coda: p-value. Per confronti, il programma dà il valore corrispondente della Normale (standardizzata); il p-value nel caso della Normale si ottiene inizialmente con la z data dalla (1) e standardizzata utilizzando la (2). Una stima del parametro, diversa può essere applicata, sia per costruire la distribuzione discreta, sia nella (1); in questa, per lo schema senza reimmissione, si utilizza la formula della var(dp) (sotto radice) riportata nella Sezione 4.1 di pag. 14.

se si dà  $\alpha = -1$ , viene utilizzato ancora come deviatore, e permette: a) di seguire l'analisi dettagliata delle elaborazioni; b) uscire dal programma. Nel caso a) si rientra in sequenza per avere i valori critici della variabile dp: dpc1 (nel testo dp') e dpc2 (nel testo dp'') o la frequenza sulla coda: p-value. Ed inoltre, l'operatore può cambiare  $\alpha$ , e specificare la direzionalità del test.

In una fase successiva, se l'operatore non vuole cambiare  $\alpha$ , ma vuole seguire le varie fasi della elaborazione, occorre dare:

3) un indice del tipo di estrazione: 0 se con reimmissione; 10, 100, 1000 o 10000 se senza reimmissione e con frazione sondata rispettivamente pari a .1, .01, .001 o .0001.

Nel caso di estrazione senza reimmissione occorre chiarire che l'indice è anche il denominatore della frazione sondata f.s. =  $n_2/N = 1/\text{indice}$ ; pertanto in questo caso, dato  $n_2$  e l'indice, rimane definita la numerosità della popolazione,  $N (= n_2 \times \text{indice})$ .

4) un indice -1 e si può successivamente dare una popolazione qualsiasi premendo Invio e dando N e NP (interi). In tal modo si offre all'operatore la possibilità di adeguarsi ad una particolare realtà a lui nota e, come di solito avviene, utilizzando l'estrazione senza reimmissione. In tal caso si ha  $P = NP/N$  e vengono escluse le due seguenti possibilità, valide per il resto.

5) l'eventuale livello di confidenza della variabile P, CONFL ( $1 - \alpha$ ), o 0 se non si vuole determinare l'intervallo; nel primo caso il programma stampa gli estremi dell'intervallo di P;

6) la P, suggerita dal programma o stabilita dall'operatore.

Durante la ricerca del p-value e nell'attuale fase, a richiesta del programma l'operatore può scegliere se fare stampare aspetti parziali delle elaborazioni dando 1 per il SI o 0 per il NO; il computer provvede ad instradare

l'operatore. Come già detto la procedura non richiede nessun calcolo, inseriti i dati dei campioni il calcolatore determina la differenza sperimentale, il valore numerico dello stimatore di P e volendo, l'intervallo di confidenza di P. Successivamente l'operatore deve solamente seguire le indicazioni e rispondere alle domande che il programma gli rivolge. Il programma può stampare anche i valori di  $\gamma_1(dp)$  e  $\gamma_2(dp)$ . In Appendice DPT sono riportati alcuni esempi di uso del programma.

Per le Tavole, si dà 1 all'inizio; poi:  $n_1, n_2$ , passo di  $n_2$ , limite di  $n_2$ , e 0 se è  $n_1 = n_2$  crescenti, o 1 per  $n_2$  crescente.

Ringrazio il Prof. Giuseppe Cicchitelli per il materiale bibliografico.

Ringrazio il Prof. Michele Grassi (vedere Sezione 3.3, pag. 11) per i numerosi colloqui avuti sui principali argomenti e con il quale purtroppo si è interrotta la collaborazione per suoi sopravvenuti impegni.

## NOTE

<sup>1)</sup> Nel test di Fisher, molto discusso e avversato in letteratura [1], [2], [3], tale vincolo limita la distribuzione campionaria ad un sottoinsieme dell'universo dei campioni.

<sup>2)</sup> Posto  $m = n_1 + n_2$ , dalla seconda delle (2),  $p = (a_1 + a_2)/m$ , e con procedimento analogo a quello seguito per la dp nelle (4) e (5), sotto  $H_0$  per lo stimatore p si ha:

$$\mu_1(p) = \pi;$$

$$\mu_2(p) = \pi(1 - \pi)/m;$$

$$\mu_3(p) = \pi(1 - \pi)(1 - 2\pi)/m^2;$$

$$\mu_4(p) = \pi(1 - \pi)[1 + 3(m - 2)\pi(1 - \pi)]/m^3;$$

$$\text{da cui: } \gamma_1(p) = (1 - 2\pi) / \sqrt{m\pi(1 - \pi)}; \quad \gamma_2(p) = [1 - 6\pi(1 - \pi)] / m\pi(1 - \pi).$$

<sup>3)</sup> Come noto, nel caso di un campione, per migliorare l'approssimazione alla Normale, per coda destra o sinistra, si può sottrarre o sommare al numeratore di z la correzione  $1/2n$ , ampiezza di un semi intervallo tra due successive determinazioni della variabile  $p = a/n$ . Nello stesso caso è nota anche la relazione:

$\chi^2 = (np - n\pi)^2 / n\pi + [n(1-p) - n(1 - \pi)]^2 / n(1 - \pi) = (np - n\pi)^2 / n(\pi(1 - \pi)) = z^2$ . Da qui, introducendo la correzione di continuità, si ha la coincidenza:  $z^2 = (|np - n\pi| - .5)^2 / n[\pi(1 - \pi)] = (|np - n\pi| - .5)^2 / n\pi + [n(1-p) - n(1 - \pi) - .5]^2 / n(1 - \pi) = \chi^2$ . O più semplicemente, eliminando i valori assoluti, nella z se  $np - n\pi > 0 < 0$ , coda destra o sinistra, si sottrae o si somma .5; ed inoltre nel  $\chi^2$  si somma o si sottrae .5 se  $n(1-p) - n(1 - \pi) > 0 < 0$ .

Ma la correzione non sembra possibile nel caso di due campioni; infatti, partendo dalla relazione:

$\chi^2 = n(a_1b_2 - b_1a_2) / n_1n_2n_a n_b = z^2$ , in cui  $z = (p_1 - p_2) / \sqrt{pq(1/n_1 + 1/n_2)}$ , con la correzione si ha:  $\chi^2 = n(|a_1b_2 - b_1a_2| - n/2)^2 / n_1n_2n_a n_b = z^2$ , in cui  $z = [dp \pm (1/2n_1 + 1/2n_2)] / \sqrt{pq(1/n_1 + 1/n_2)}$  (+ per coda sinistra, - per coda destra). Tale correzione non può essere utilizzata. Infatti anche nel caso più semplice di  $n_1 = n_2 = n$ , in cui gli intervalli tra un valore ed il successivo della variabile dp siano costanti, la correzione risulta  $1/2n + 1/2n = 1/n$ , che non corrisponde ad un semi intervallo. Ad es. per  $n_1 = n_2 = 4$  i valori della variabile dp risultano: -1, -.75, -.5, -.25, 0, .25, .5, .75, 1, con intervalli di .25. L'istogramma presenta i successivi rettangoli aventi la stessa base di ampiezza .25; graficamente la correzione dovrebbe corrispondere ad un semi intervallo di ampiezza 0.125, mentre la correzione sopra vista vale  $1/8 + 1/8 = .25$ . Per  $n_1 \neq n_2$  la situazione è ancora peggiore perché gli intervalli tra un valore e l'altro possono non essere costanti, e ciò si verifica soprattutto sulle code delle distribuzioni in prossimità delle zone di rifiuto.

<sup>4)</sup> Come noto, un valore critico nel discontinuo è il valore che delimita la zona di rifiuto avente la frequenza più vicina e non superiore ad  $\alpha$  o ad  $\alpha/2$ ; per numerosità molto piccole e quando i valori di  $\pi$  sono prossimi a 0 o ad 1 a volte non è possibile determinarlo perché la frequenza del primo o dell'ultimo valore della variabile dp è maggiore di  $\alpha$  o  $\alpha/2$ ; inoltre per valori di  $\pi$  prossimi a .5 e contigui si può trovare lo stesso valore critico; quindi la variabilità e il dettaglio dei valori critici è legata al dettaglio dei valori di  $\pi$  considerati.

5) Se i campioni hanno numerosità diverse, non conviene usare valori interi come nella Table 4 (in Appendice 4) del precedente lavoro [14] ( $n_1 = n_2 = n$ , quindi  $n \times dp$  intero). In quel caso, essendo  $a_1$  e  $a_2$  interi, sia il valore sperimentale della differenza  $d = a_1 - a_2$ , sia i valori critici potevano essere espressi in valori interi. Ma nel caso attuale con  $n_1 \neq n_2$ , si ha:  $dp = p_1 - p_2 = a_1/n_1 - a_2/n_2 = (n_2 a_1 - n_1 a_2)/n_1 n_2$ , ed il corrispondente intero dovrebbero risultare dall'espressione  $n_1 n_2 dp = n_2 a_1 - n_1 a_2$ , calcolato dal programma, ma che complica notevolmente la interpretazione da parte dell'operatore. Pertanto è più semplice tabulare i valori critici della differenza  $dp = p_1 - p_2$ , il cui valore sperimentale lo stesso operatore può calcolare facilmente.

6) Come nel continuo, se un risultato sperimentale è uguale al valore critico,  $H_0$  viene rifiutata.

7) La potenza può essere calcolata direttamente partendo dallo stesso valore critico  $dp^*$ , incluso (coda destra):

$$1 - \beta = \sum_{dp=dp^*}^1 f(dp | \delta, \pi_1)$$

8) Commento strettamente statistico - campionario. Nei due casi le distribuzioni hanno scarsissima variabilità, elevata curtosi e asimmetria contraria; nel primo l'asimmetria positiva (coda destra più lunga) concentra le frequenze più elevate verso la coda a sinistra della moda della distribuzione: il valore della  $dp_s$  cumula una frequenza  $> \alpha$ ; nel secondo l'asimmetria negativa concentra le frequenze meno elevate verso la coda a sinistra: il valore della  $dp_s$  cumula una frequenza  $< \alpha$ .

9) Come noto, analogamente alla differenza tra due medie, la distribuzione si ottiene partendo da due popolazioni di parametri  $\pi_1$  e  $\pi_2$ ; dalla prima viene estratto un campione di numerosità  $n_1$ , dalla seconda un campione di numerosità  $n_2$ , con numeri di successi rispettivamente  $a_1$  e  $a_2$  che danno luogo alle proporzioni  $p_1$  e  $p_2$ .

La distribuzione ha media  $E(dp) = \pi_1 - \pi_2$  e, se  $N_1 \neq N_2$ ,

$$\text{var}(dp) = (N_1 - n_1) \pi_1 (1 - \pi_1) / (N_1 - 1) n_1 + (N_2 - n_2) \pi_2 (1 - \pi_2) / (N_2 - 1) n_2,$$

somma delle due varianze. Sotto  $H_0$ ,  $\pi_1 = \pi_2 = \pi$ , e considerato per semplicità ed economia di elaborazioni

$N_1 = N_2 = N$  si ottiene l'espressione del testo. Più in generale, se vale  $H_0$ , si dice che i due campioni sono estratti da due popolazioni aventi la stessa  $\pi$  (con  $N_1 = 0 \neq N_2$ ) o anche dalla stessa popolazione (se  $N_1 = N_2 = N$ ).

10) Mediante la (11) è facile determinare i seguenti momenti della variabile  $dp$ :

$$E\left(\frac{a_1}{n_1}\right) - E\left(\frac{a_2}{n_2}\right) = \pi - \pi = 0; \quad E\left(\frac{a_1}{n_1} \frac{a_2}{n_2}\right) = \frac{N\pi^2 - \pi}{N-1} = \pi^2 - \frac{\pi(1-\pi)}{N-1}; \quad \text{cov}\left(\frac{a_1}{n_1}, \frac{a_2}{n_2}\right) = -\frac{\pi(1-\pi)}{N-1}.$$

## BIBLIOGRAFIA

- [1] Ballatori E., Sui test statistici per il confronto fra due frequenze in tabelle 2x2, Metron, Vol. XL - N. 3-4 1982, p 167-169.
- [2] Ballatori E., Generalizing Fisher's exact test, Metron, Vol. XLV-N 3-4 1987, p. 99.
- [3] Berkson J., In dispraise of the exact test. Do the marginal totals of the 2x2 tables contain relevant information respecting the table proportions? J.S.P.I.1978,2, p.27-44.
- [4] Chan I. S. F., Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies, 1998, Statistics in Medicine, 17, 1403-1413.
- [5] Chan I. S. F., Zhang Z., Test-based exact confidence intervals for the difference of two binomial proportions, 1999, Biometrics 55, 1202 - 1209.
- [6] Cicchitelli G., Herzel A., G.E. Montanari, Il campionamento statistico, il Mulino, Bologna, 1992, p 57,58.

- [7] Clopper C. J. & Pearson E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, XXVI, 1934, p. 404-413.
- [8] Cochran W. G., *Sampling Techniques*, 1977, John Wiley & Sons, N. Y., London, p. 21 e 259.
- [9] Cramer H., *Mathematical Methods of Statistics*, Princeton, Princeton University Press, 1963, p. 511-515.
- [10] De Lucia L., *Problemi di tecnica campionaria nelle analisi di mercato*, Giuffrè, Milano, 1958 p. 9.
- [11] Herzel A., *Sui momenti centrali dei campioni estratti con e senza ripetizione*. Biblioteca del "Metron" Serie C: note e commenti. Istituto di Statistica, Roma, 1961.
- [12] Gori F., *Il coefficiente di correlazione r nel caso di caratteri dicotomici e la differenza tra proporzioni, negli universi con e senza ripetizione per piccoli campioni*. Serie A– Ricerche, 1989, N. 20. Dipartimento di Statistica, Probabilità e Statistiche Applicate.
- [13] Gori F., *Elementi di Statistica, Parte II*, Edizioni Ifrep Roma, 1993, p. 304, 305
- [14] Gori F.-Leone L., *The difference between proportions in small independent samples in sampling with and without replacement*, T.P.M. Vol. 9 N. 3 Padova, Autunno 2002, p.77-92.
- [15] Gori F.-Leone L., *Difference between proportions and correlations in small paired groups sampled with and without replacement*, T.P.M. Vol. 11 No.1 Primavera 2004, p.23-46.
- [16] Mehrotra D. V., Chan I.S.F., Berger R. L., *A Cautionary Note on Exact Unconditional Inference for a Difference between Two Independent Binomial Proportions*, 2003, *Biometrics* 59, 441 – 450
- [17] Phipps, M. C., & Byron, P. M. , *A filter for confidence interval P-values*, 2007, *Computational Statistics & Data Analysis*, 51 (12), 6435-6446.
- [18] Sarndal C. E., Swensson B., Wretman J.: *Model Assisted Survey Sampling*, 1997 N. Y., p. 30, 31 e 66.
- [19] Spiegel M. R., *Statistica, Collana Schaum, Etas Libri*, 1976, Milano, p.188.

## APPENDICE 1

Per  $\pi_1 = .15, .5$  e  $.85$ , i tre prospetti riportano i valori di  $\delta | \pi_1$ , del parametro  $\pi_2$  e della potenza per un test monolaterale destro,  $\alpha = 0.025$ ;  $n_1 = 5, n_2 = 10$ , estrazione con ripetizione.

Prospetto 1 Sotto $H_0$ $dp'' = .5$ , con n° d'ordine: 16			Prospetto 2 Sotto $H_0$ $dp'' = .6$ , con n° d'ordine: 17			Prospetto 3 Sotto $H_0$ $dp'' = .4$ , con n° d'ordine: 15		
$\delta   .15$	$\pi_2$	$1-\beta$	$\delta   .5$	$\pi_2$	$1-\beta$	$\delta   .85$	$\pi_2$	$1-\beta$
-.8	.95	$\cong 0$	-.45	.95	$\cong 0$	-.1	.95	0.0005
-.75	.9	$\cong 0$	-.4	.9	$\cong 0$	-.05	.9	0.0057
-.7	.85	$\cong 0$	-.35	.85	$\cong 0$	<b>0</b>	<b>.85</b>	<b>0.0227</b>
-.65	.8	$\cong 0$	-.3	.8	0.0002	.05	.8	0.0561
-.6	.75	$\cong 0$	-.25	.75	0.0007	.1	.75	0.1072
-.55	.7	$\cong 0$	-.2	.7	0.0017	.15	.7	0.1742
-.5	.65	$\cong 0$	-.15	.65	0.0037	.2	.65	0.2536
-.45	.6	0.0002	-.1	.6	0.0071	.25	.6	0.3409
-.4	.55	0.0004	-.05	.55	0.0126	.3	.55	0.4319
-.35	.5	0.0007	<b>0</b>	<b>.5</b>	<b>0.0206</b>	.35	.5	0.5226
-.3	.45	0.0012	.05	.45	0.0321	.4	.45	0.6097
-.25	.4	0.0020	.1	.4	0.0478	.45	.4	0.6905
-.2	.35	0.0033	.15	.35	0.0686	.5	.35	0.7628
-.15	.3	0.0051	.2	.3	0.0952	.55	.3	0.8253
-.1	.25	0.0077	.25	.25	0.1285	.6	.25	0.8771
-.05	.2	0.0111	.3	.2	0.1697	.65	.2	0.9183
<b>0</b>	<b>.15</b>	<b>0.0154</b>	.35	.15	0.2206	.7	.15	0.9494
.05	.1	0.0201	.4	.1	0.2854	.75	.1	0.9715
.1	.05	0.0245	.45	.05	0.3728	.8	.05	0.9864

## APPENDICE 2

Per  $\pi_1 = .15, .5$  e  $.85$ , i tre prospetti riportano i valori di  $\delta | \pi_1$ , del parametro  $\pi_2$  e della potenza, per un test monolaterale sinistro,  $\alpha = 0.025$ ;  $n_1 = 5, n_2 = 10$ , estrazione con ripetizione. I n° d'ordine sono: quello relativo a  $dp'$  e l'ultimo della distribuzione.

Prospetto 1 Sotto $H_0$ $dp' = -.4$ n° d'ordine 7 e 21			Prospetto 2 Sotto $H_0$ $dp' = -.6$ n° d'ordine 5 e 21			Prospetto 3 Sotto $H_0$ $dp' = -.5$ n° d'ordine 6 e 21		
$\delta   .15$	$\pi_2$	$1 - \beta$	$\delta   .5$	$\pi_2$	$1 - \beta$	$\delta   .85$	$\pi_2$	$1 - \beta$
-.8	.95	0.9864	-.45	.95	0.3728	-.1	.95	0.0245
-.75	.9	0.9715	-.4	.9	0.2854	-.05	.9	0.0201
-.7	.85	0.9494	-.35	.85	0.2206	<b>0</b>	<b>.85</b>	<b>0.0154</b>
-.65	.8	0.9183	-.3	.8	0.1697	.05	.8	0.0111
-.6	.75	0.8771	-.25	.75	0.1285	.1	.75	0.0077
-.55	.7	0.8253	-.2	.7	0.0952	.15	.7	0.0051
-.5	.65	0.7628	-.15	.65	0.0686	.2	.65	0.0033
-.45	.6	0.6905	-.1	.6	0.0478	.25	.6	0.0020
-.4	.55	0.6097	-.05	.55	0.0321	.3	.55	0.0012
-.35	.5	0.5226	<b>0</b>	<b>.5</b>	<b>0.0206</b>	.35	.5	0.0007
-.3	.45	0.4319	.05	.45	0.0126	.4	.45	0.0004
-.25	.4	0.3409	.1	.4	0.0071	.45	.4	0.0002
-.2	.35	0.2536	.15	.35	0.0037	.5	.35	0.0000
-.15	.3	0.1742	.2	.3	0.0017	.55	.3	0.0000
-.1	.25	0.1072	.25	.25	0.0007	.6	.25	0.0000
-.05	.2	0.0561	.3	.2	0.0002	.65	.2	0.0000
<b>0</b>	<b>.15</b>	<b>0.0227</b>	.35	.15	0.0000	.7	.15	0.0000
.05	.1	0.0057	.4	.1	0.0000	.75	.1	0.0000
.1	.05	0.0005	.45	.05	0.0000	.8	.05	0.0000

## APPENDICE 3

Per  $\pi_1 = .15, .5$  e  $.85$ , i tre prospetti riportano i valori di  $\delta | \pi_1$ , del parametro  $\pi_2$  e della potenza, per un test bilaterale,  $\alpha = 0.05$ ;  $n_1 = 5, n_2 = 10$ , estrazione con ripetizione.

Prospetto 1 Sotto $H_0$ , $dp' = -.4, dp'' = .5$ con n° d'ordine 7 e 16			Prospetto 2 Sotto $H_0$ , $dp' = -.6, dp'' = .6$ con n° d'ordine 5 e 17			Prospetto 3 Sotto $H_0$ , $dp' = -.5, dp'' = .4$ con n° d'ordine 6 e 15		
$\delta   .15$	$\pi_2$	$1 - \beta$	$\delta   .5$	$\pi_2$	$1 - \beta$	$\delta   .85$	$\pi_2$	$1 - \beta$
-.8	.95	0.9864	-.45	.95	0.3728	-.1	.95	0.0250
-.75	.9	0.9715	-.4	.9	0.2854	-.05	.9	0.0259
-.7	.85	0.9494	-.35	.85	0.2207	<b>0</b>	<b>.85</b>	<b>0.0381</b>
-.65	.8	0.9183	-.3	.8	0.1699	.05	.8	0.0673
-.6	.75	0.8771	-.25	.75	0.1292	.1	.75	0.1149
-.55	.7	0.8253	-.2	.7	0.0969	.15	.7	0.1793
-.5	.65	0.7629	-.15	.65	0.0723	.2	.65	0.2568
-.45	.6	0.6907	-.1	.6	0.0550	.25	.6	0.3429
-.4	.55	0.6101	-.05	.55	0.0447	.3	.55	0.4331
-.35	.5	0.5233	<b>0</b>	<b>.5</b>	<b>0.0413</b>	.35	.5	0.5233
-.3	.45	0.4331	.05	.45	0.0447	.4	.45	0.6101
-.25	.4	0.3429	.1	.4	0.0550	.45	.4	0.6907
-.2	.35	0.2568	.15	.35	0.0723	.5	.35	0.7629
-.15	.3	0.1793	.2	.3	0.0969	.55	.3	0.8253
-.1	.25	0.1149	.25	.25	0.1292	.6	.25	0.8771
-.05	.2	0.0673	.3	.2	0.1699	.65	.2	0.9183
<b>0</b>	<b>.15</b>	<b>0.0381</b>	.35	.15	0.2207	.7	.15	0.9494
.05	.1	0.0259	.4	.1	0.2854	.75	.1	0.9715
.1	.05	0.0250	.45	.05	0.3728	.8	.05	0.9864

## APPENDICE 4

Table 4. Critical values <sup>5)</sup> of  $\underline{d}$  for a one-tailed test for independent samples drawn without and with replacement for sample sizes  $\underline{n} = 3, (1), 30$ ,  $\mathcal{T} = .05, (.05), .5$  and for  $\alpha = .05$  (first row) and  $\alpha = .025$  (second row). In the cells with more than one value of  $\underline{d}$ , the first value refer to sampling without replacement for  $f = .1$ , the second again to the sampling without replacement but for  $f = .01$ , and so on; the last value refer always *also* to sampling with replacement. For example if  $\mathcal{T} = .5$ ,  $\alpha = .025$  and  $\underline{n} = 16$ , we have:  $\underline{d} = 6$  for  $f = .1, .01, .001$  and  $\underline{d} = 7$  for  $f = .0001$  and with replacement; but if  $\underline{n} = 17$ :  $\underline{d} = 6$  for  $f = .1$  and  $\underline{d} = 7$  for  $f = .01, .001, .0001$  and with replacement. When only one value appears in the cell, it is applicable to both sampling methods. For  $\mathcal{T} = .05, .15, .25, .35, .45$ , when  $\underline{n}$  is odd the first  $\underline{d}$  value for sampling without replacement refers to  $f = .01$ .

$n \backslash \pi$	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5
3	2	2	2	2,3	3	3	3	3	3	3
			3	3						
4	2	2	3	3	3	3	3	3	3	3
		3					3,4	4	4	4
5	2	2,3	3	3	3	3	3	3,4	4	3,4
		3			4	4	4	4		
6	2	3	3	3	3	3,4	4	4	4	4
				4	4	4				
7	2	3	3	3	4	4	4	4	4	4
	3		4	4				4,5	5	4,5
8	2	3	3	3,4	4	4	4	4	4	4
	3		4	4		4,5	5	5	5	5
9	2	3	3	4	4	4	4	4	4	4
	3		4		5	5	5	5	5	5
10	2,3	3	3,4	4	4	4	4	4,5	4,5	4,5
	3	3,4	4		5	5	5	5	5	5
11	3	3	4	4	4	4,4,5	5	5	5	5
		4		4,5	5	5			6	5,6
12	3	3	4	4	4	4,5	5	5	5	5
		4		5	5	5	5,6	5,6	6	6
13	3	3	4	4	5	5	5	5	5	5
		4	5	5		5,6	6	6	6	6
14	3	3,4	4	4	5	5	5	5	5	5
		4	4,5	5		5,6	6	6	6	6
15	3	4	4	4,5	5	5	5	5	5	5
			5	5	6	6	6	6	6	6
16	3	4	4	5	5	5	5	5,6	5,6	5,6
			5	6	6	6	6	6	6	6,6,6,7
17	3	4	4	5	5	5	6	5,6	6	6
			5	5,5,6	6	6		6,7	7	6,7
18	3	4	4	5	5	5	5,6	6	6	6
	3,4	4,4,5	5	5,6	6	6	6,7	6,7	7	7
19	3	4	5	5	5	5,6	6	6	6	6
	4	4,5		6	6	6,6,6,7	7	7	7	7
20	3	4	5	5	5	6	6	6	6	6
	4	5		6	6	6,7	7	7	7	7
21	3	4	5	5	6	6	6	6	6	6
	4	5	5,6	6		7	7	7	7	7
22	3	4	5	5	5,6	6	6	6	6	6
	4	5	5,6	6	7	7	7	7	7	7
23	3	4	5	5	6	6	6	6	7	6,7
	4	5	6	6	7	7	7	7	8	7,8
24	3	4	5	5,6	6	6	6	6,7	6,7	6,7
	4	5	6	6	7	7	7	7,8	7,8	7,8
25	3	4	5	5,6	6	6	7	6,7	7	7
	4	5	6	6,6,7	7	7	8	7,8	8	8
26	3,4	4,5	5	5,6	6	6	6,7	7	7	7
	4	5	6	6,7	7	7	7,8	8	8	8
27	4	4,5	5	6	6	6,6,7	7	7	7	7
		5	6	6,7	7	7,8	8	8	8	8

28	4	4,5	5	6	6	6,7	7	7	7	7
		5	6	7	7	7,8	8	8	8	8
29	4	5	5	6	6	6,7	7	7	7	7
			6	7	7	7,8	8	8	8	8
30	4	5	5,6	6	6	7	7	7	7	7
		5,6	6	7	8	8	8	8	8,9	8,9

## APPENDICE DPT

I Esempio. Ipotesi bidirezionale con  $\alpha=0.05$ , due campioni di ampiezza  $n_1=15$  e  $n_2=20$ , numeri di successi  $a_1=10$  e  $a_2=10$ .

Si inizia cliccando sull'icona DPT.EXE.

TEST DIFFERENZA TRA PROPORZIONI

DUE CAMPIONI INDIPENDENTI, AMPIEZZA  $\leq 30$

Input: dopo i due punti dare valori spazati e premere INVIO.

Quando il programma si ferma premere INVIO

TEST=0, TAVOLE=1: **0**

Numerosità dei 2 campioni,  $n_1 \leq n_2$ : **15 20**

n° successi  $a_1$  e  $a_2$  nei 2 campioni

$a_1 \leq n_1, a_2 \leq n_2$ : **10 10**

Proporzioni  $p_1$  e  $p_2$  nei 2 campioni

0.66666666666666663 0.50000000000000000

Differenza sperimentale  $dps=p_1-p_2=$  0.16666666666666666

Stima di P,  $stp=(n_1*p_1+n_2*p_2)/(n_1+n_2)=$  0.57142860

In questa fase la estrazione è con reimmissione

Dare  $\alpha$ ,  $0 < \alpha < .5$ , prob. errori di I specie:

inoltre,

$se=0$ , f.c. sulla coda verso dps,

$se=-1$ , analisi elaborati o STOP: **.05**

Per una ipotesi monodirezionale sinistra dare -1,

" " " bidirezionale dare 0,

" " " monodirezionale destra dare 1: **0**

CODA SINISTRA

$n_1=$  15  $n_2=$  20

Alfa = 5.00000007E-02 Prob. rifiuto,coda sinistra= 2.500000000000000500E-02

Valore sperimentale dps

0.16666666666666666

VALORE CRITICO  $dpc1$  f.c.( $dpc1$ )

-0.34999999999999998 2.19245057075477222E-02

CODA DESTRA

$n_1=$  15  $n_2=$  20

Alfa = 5.00000007E-02 Prob. rifiuto,coda destra= 2.500000000000000500E-02

Valore sperimentale dps

0.16666666666666666

VALORE CRITICO  $dpc2$  f.rtc.( $dpc2$ )

0.34999999999999998 2.09096126004583993E-02

Cambiare  $\alpha$ ? SI=1, NO=0: **0**

Seguire elaborazioni e cambiare tipo di estraz.?

SI=1, NO=0 e STOP: **0**

COMMENTO: Per  $P=0.5714$  la  $dps$  è  $<$  al valore critico di destra: il test non è significativo.

Il Esempio. Due campioni di ampiezza  $n_1=12$  e  $n_2=18$ , numero di successi  $a_1=9$  e  $a_2=9$ , si vuole determinare il p-value della  $dps$ .

TEST DIFFERENZA TRA PROPORZIONI

DUE CAMPIONI INDIPENDENTI, AMPIEZZA  $\leq 30$

Input: dopo i due punti dare valori spazati e premere INVIO.

Quando il programma si ferma premere INVIO

TEST=0, TAVOLE=1: **0**

Numerosità dei 2 campioni,  $n_1 \leq n_2$ : **12 18**

n° successi  $a_1$  e  $a_2$  nei 2 campioni

$a_1 \leq n_1, a_2 \leq n_2$ : **9 9**

Proporzioni  $p_1$  e  $p_2$  nei 2 campioni

0.75000000000000000 0.50000000000000000

Differenza sperimentale  $dps=p_1-p_2=$  0.25000000000000000

Stima di P,  $stp=(n_1*p_1+n_2*p_2)/(n_1+n_2)=$  0.60000002

In questa fase la estrazione è con reimmissione

Dare  $\alpha$ ,  $0 < \alpha < .5$ , prob. errori di I specie:

inoltre,

$se=0$ , f.c. sulla coda verso dps,

$se=-1$ , analisi elaborati o STOP: **0**

Stampe parz.? SI=1, NO=0: **0**

Coda destra

dps

f.rtc.

0.2500000000000000 9.80885575669502435E-02  
z = 1.36930639376291530 Normale 8.54514683626940164E-02  
Dare alfa, 0 < alfa < .5, prob. errori di I specie:  
inoltre,  
se=0, f.c. sulla coda verso dps,  
se=-1, analisi elaborati o STOP: **-1**  
Seguire elaborazioni e cambiare tipo di estraz.?  
SI=1, NO=0 e STOP: **0**  
COMMENTO: Per P = .6, alla dps corrisponde un p-value = .098. La Normale standardizzata dà .085.

III Esempio. Si considerano i tre possibili input di una popolazione con parametro .6. IL primo nel bernoulliano, gli altri due nell'esautivo.

TEST DIFFERENZA TRA PROPORZIONI  
DUE CAMPIONI INDIPENDENTI, AMPIEZZA <= 30  
Input: dopo i due punti dare valori spazati e premere INVIO.  
Quando il programma si ferma premere INVIO  
TEST=0, TAVOLE=1: **0**  
Numerosità dei 2 campioni, n1 <= n2: **5 5**  
n° successi a1 e a2 nei 2 campioni  
a1 <= n1, a2 <= n2: **5 1**  
Proporzioni p1 e p2 nei 2 campioni  
1.0000000000000000 0.2000000000000000  
Differenza sperimentale dps=p1-p2= 0.8000000000000004  
Stima di P, stp=(n1\*p1+n2\*p2)/(n1+n2)= 0.60000000  
In questa fase la estrazione è con reimmissione  
Dare alfa, 0 < alfa < .5, prob. errori di I specie:  
inoltre,  
se=0, f.c. sulla coda verso dps,  
se=-1, analisi elaborati o STOP: **.05**  
Per una ipotesi monodirezionale sinistra dare -1,  
" " " bidirezionale dare 0,  
" " " monodirezionale destra dare 1: **1**  
CODA DESTRA  
n1= 5 n2= 5  
Alfa = 5.00000007E-02 Prob. rifiuto,coda destra= 5.00000000000000999E-02  
Valore sperimentale dps  
0.8000000000000004  
VALORE CRITICO dpc2 f.rtc.(dpc2)  
0.8000000000000004 9.42243840000000186E-03  
VALORE CRITICO (intero) della variabile nxdp =n1 x 0.8000000000000004  
4  
Cambiare alfa? SI=1, NO=0: **0**  
Seguire elaborazioni e cambiare tipo di estraz.?  
SI=1, NO=0 e STOP: **1**  
Estrazione con reimmissione dare 0  
Estrazione senza reim. e per frazione  
sondata .1, o .01, o .001, o .0001  
dare 10, o 100, o 1000, o 10000  
Popol. da input,dare -1,INVIO,N,NP(<N)  
: **10**  
Intervallo di confidenza di P ? NO=0, SI,  
Dare Livello di Confidenza, CONFL: **0**  
Dare il valore di P  
o = alla stima 0.5999999999999998  
o altro valore di interesse:  
con numero di cifre decimali < 2: **.6**  
Popolaz. sondata  
N Np N-Np  
50 30 20  
Np/N= 0.5999999999999998  
Se Np/N=P data, OK, dare 1; altrimenti 0: **1**  
Cambiare fraz. sond.? SI=1, NO=0: **0**  
Stampe elabor. variab. dp? SI=1, NO=0: **0**  
Dare alfa, 0 < alfa < .5, prob. errori di I specie:  
inoltre,  
se=0, f.c. sulla coda verso dps,  
se=-1, analisi elaborati o STOP: **.05**  
Per una ipotesi monodirezionale sinistra dare -1,  
" " " bidirezionale dare 0,  
" " " monodirezionale destra dare 1: **1**  
CODA DESTRA  
n1= 5 n2= 5  
Alfa = 5.00000007E-02 Prob. rifiuto,coda destra= 5.00000000000000999E-02  
Valore sperimentale dps  
0.8000000000000004  
VALORE CRITICO dpc2 f.rtc.(dpc2)  
0.5999999999999998 4.31519352501681316E-02

VALORE CRITICO (intero) della variabile  $nxdp = n1 \times \frac{3}{5}$  0.5999999999999998

Cambiare alfa? SI=1, NO=0: **0**

Seguire elaborazioni e cambiare tipo di estraz.?

SI=1, NO=0 e STOP: **1**

Estrazione con reimmissione dare 0

Estrazione senza reim. e per frazione

sondata .1, o .01, o .001, o .0001

dare 10, o 100, o 1000, o 10000

Popol. da input,dare -1,INVIO,N,NP(<N)

: **-1**

**50 30**

NP,NQ,N,NP/N

30 20 50 0.5999999999999998

Stampe elabor. variab. dp? SI=1, NO=0: **0**

Dare alfa,  $0 < \alpha < .5$ , prob. errori di I specie:

inoltre,

se=0, f.c. sulla coda verso dps,

se=-1, analisi elaborati o STOP: **.05**

Per una ipotesi monodirezionale sinistra dare -1,

" " " bidirezionale dare 0,

" " " monodirezionale destra dare 1: **1**

CODA DESTRA

n1= 5 n2= 5

Alfa = 5.00000007E-02 Prob. rifiuto,coda destra= 5.00000000000000999E-02

Valore sperimentale dps

0.80000000000000004

VALORE CRITICO dpc2 f.rtc.(dpc2)

0.5999999999999998 4.31519352501681316E-02

VALORE CRITICO (intero) della variabile  $nxdp = n1 \times \frac{3}{5}$  0.5999999999999998

Cambiare alfa? SI=1, NO=0: **0**

Seguire elaborazioni e cambiare tipo di estraz.?

SI=1, NO=0 e STOP: **0**

COMMENTO. Nel primo caso, estrazione con reimmissione, il valore critico è .8, negli altri due, senza reimmissione (minor variabilità), è .6.

IV Esempio Determinazione dell'intervallo di confidenza.

TEST DIFFERENZA TRA PROPORZIONI

DUE CAMPIONI INDIPENDENTI,AMPIEZZA  $\leq 30$

Input: dopo i due punti dare valori spaziosi e

premere INVIO.

Quando il programma si ferma premere INVIO

TEST=0, TAVOLE=1: **0**

Numerosità dei 2 campioni,  $n1 \leq n2$ : **12 16**

n° successi a1 e a2 nei 2 campioni

a1  $\leq n1$ , a2  $\leq n2$ : **3 10**

Proporzioni p1 e p2 nei 2 campioni

0.2500000000000000 0.6250000000000000

Differenza sperimentale  $dps=p1-p2=$  -0.3750000000000000

Stima di P,  $stp=(n1*p1+n2*p2)/(n1+n2)=$  0.46428570

In questa fase la estrazione è con reimmissione

Dare alfa,  $0 < \alpha < .5$ , prob. errori di I specie:

inoltre,

se=0, f.c. sulla coda verso dps,

se=-1, analisi elaborati o STOP: **-1**

Seguire elaborazioni e cambiare tipo di estraz.?

SI=1, NO=0 e STOP: **1**

Estrazione con reimmissione dare 0

Estrazione senza reim. e per frazione

sondata .1, o .01, o .001, o .0001

dare 10, o 100, o 1000, o 10000

Popol. da input,dare -1,INVIO,N,NP(<N)

: **0**

Intervallo di confidenza di P ? NO=0, SI,

Dare Livello di Confidenza, CONFL: **.9**

Stampe costruzione intervallo,

tutte o in parte=1, niente=0: **0**

$(1-CONFL)/2=$  4.999999999999889E-02

Intervallo di confidenza

P1= 0.30000001192092895 P2= 0.65000003576278686

Dare il valore di P

o = alla stima 0.46428571428571430

o altro valore di interesse: proseguire in maniera analoga al precedente Esempio