Tipo variabile Variabile X	Analisi Monovariata (una variabile) Indici di dimensione (o di posizione) oppure Misure di tendenza centrale	Analisi Monovariata (una variabile) Indici di variabilità (o di dispersione)	Analisi Bivariata (2 variabili) Variabile Y (la relazione tra variab. nominali si dice Associazione; tra variab. ordinali Cograduazione; tra variab. cardinali Correlazione) Nominale	Cardinale
Nominale	**Moda** Mo La modalità (o le modalità) con frequenza maggiore	**Indice di omogeneità** O $p_1^2+p_2^2+\ldots+p_k^2 = \sum_{i=1}^{k} p_i^2$ p_i = proporzioni relativizzate ad 1; k = nr. modalità variabile Varia da un min di 1/k ad un max di 1 **Indice di eterogeneità** E 1 - (indice di omogeneità: O) = $1 - \sum_{i=1}^{k} p_i^2$ O = indice di omogeneità Complemento ad 1 di O **Indice di omogeneità relativa** O_{rel} (k*O - 1)/(k - 1) k = nr. modalità variabile Varia da un min di 0 ad un max di 1 **Indice di eterogeneità relativa** E_{rel} $1 - O_{rel}$ = 1 - [(k*O - 1)/(k - 1)] = [k/(k - 1)]*E O_{rel} = indice di omogeneità relativa Complemento ad 1 di O_{rel}	**Tavola di contingenza** E' una tabella o matrice a doppia entrata, ottenuta riportando sui lati le modalità delle 2 variabili e nelle celle le frequenze corrispondenti f_{ij} f_{ij} = frequenza osservata alla riga i ed alla colonna j Le frequenze possono essere assolute, percentualizzate e/o cumulate; la regola vuole che si percentualizzi all'interno della variabile "indipendente" **Indice di differenza percentuale** Si calcola la differenza fra 2 modalità per tutta la riga (o colonna) in modo da coglierne meglio l'andamento **Chi quadrato (tabulato)** χ^2 $\sum[(f_o - f_e)^2/f_e]$ f_o = frequenza osservata; f_e = freq. attesa (nell'ipotesi H_0) = (freq.marginale di riga*freq.marg.di colonna)/freq. totale di tabella Gradi di libertà (g.l.) = (nr. righe - 1) * (nr. colonne - 1); si usa per N>100 e per celle con freq>= 5; se χ^2 calcolato è maggiore di χ^2 tabulato si respinge l'ipotesi nulla H_0 (di indipendenza fra variabili) e si accetta H_1 di esistenza di una relazione **Chi quadrato di una tabella 2 x 2** χ^2 $N*(a*d - b*c)^2 / [(a+b)*(c+d)*(a+c)*(b+d)]$ a; b; c; d = frequenze assolute della tabella 2 x 2 **Phi** Φ $\sqrt{(\chi^2/N)}$ Varia da un min di -1 ad un max di +1 **Phi di una tabella 2 x 2** Φ $(a*d - b*c) / \sqrt{[(a+b)*(c+d)*(a+c)*(b+d)]}$ Varia da un min di -1 ad un max di +1; nel caso della tabella 2 x 2 (o comunque con almeno una variab. di 2) si ha Φ = V (di Cramer) = r (di Pearson) **Differenza fra proporzioni (tabella 2 x 2)** E' la semplice differenza fra le proporzioni di una colonna Varia da un min di -1 ad un max di +1; se è uguale a 0 non c'è relazione tra le variabili; la differenza dell'altra colonna ha lo stesso valore assoluto ma è di segno opposto; questo indice coincide con il Coefficiente di Regressione b **Q di Yule (tabella 2 x 2)** Q (a*d - b*c) / (a*d + b*c) Varia, come Φ, da un min di -1 ad un max di +1 **V di Cramer** V $\sqrt{[\chi^2/(N*(k-1)]}$ k = minore tra nr. righe e nr. colonne; N*(k -1) = valore max di χ^2 Varia da un min di 0 (indipendenza) ad un max di 1 (associazione perfetta) **C di Pearson o Coefficiente di Contingenza** C $\sqrt{[\chi^2/(\chi^2+ N)]}$ **Odds Ratio (tabella 2 x 2: variabile dicotomica)** ω_1/ω_2 (a*d) / (b*c) $\omega = p/(1 - p)$; $p = \omega/(1 + \omega)$ p = probabilità favorevole; ω = odds o rapporto di probabilità; sono usati nel gioco d'azzardo, dove ω indica il rapporto tra vincita e cifra scommessa	**Analisi della Varianza** Anova Devianza totale: $_i\sum{}_j\sum(x_{ij} - M_{..})^2$ = Devianza non spiegata: $_i\sum{}_j\sum(x_{ij} - M_{.j})^2$ + Devianza spiegata: $_i\sum{}_j\sum(M_{.j} - M_{..})^2$ La Dev. non spiegata (o Interna) è la parte di Dev. totale non determinata dalla variab. nominale (è uguale a 0 se la relazione è perfetta): la Dev. spiegata (o Esterna) è la parte di Dev. totale determinata dalla variab. nominale (è uguale a 0 se non c'è relazione) Gradi di libertà (g.l., con k = nr. modalità variab. nominale) = Dev. totale: (N -1); Dev. Interna o non spiegata: (N - k); Dev. esterna o spiegata: (k -1) **Rapporto F (tabulato)** F Stima Esterna: $[_i\sum{}_j\sum(M_{.j} - M_{..})^2 / (k - 1)]$ / Stima Interna: $[_i\sum{}_j\sum(x_{ij} - M_{.j})^2 / (N - k)]$ Rapporto (Stima Esterna/Stima Interna); Stima Esterna = (Dev. Esterna/g.l. Esterni); Stima Interna = (Dev Interna/g.l. Interni) Se F calcolato è maggiore di F tabulato si respinge l'ipotesi nulla H_0 e si accetta H_1 di esistenza di una relazione **Rapporto di Correlazione (forza dell'Anova: Eta quadrato)** η^2 Devianza spiegata: $_i\sum{}_j\sum(M_{.j} - M_{..})^2$ / Devianza totale: $_i\sum{}_j\sum(x_{ij} - M_{..})^2$ Rapporto (Dev. spiegata/Dev. totale) Varia da un min di 0 ad un max di 1; si considera rilevante $\eta^2 > 0,10$
Ordinale	**Mediana** Me Il valore centrale della serie, da calcolare dopo aver messo in ordine crescente tutti i valori N = nr. dei casi Se N è dispari si ha 1 mediana; se N è pari se ne hanno 2; nel caso di variabili cardinali la mediana è la media dei 2 valori centrali, cioè quelli posizionati a: N/2 e N/2 + 1 **Quantili** $Q_{(x)}$ I valori situati ad una certa posizione, da calcolare dopo aver messo in ordine crescente tutti i valori Terzili: Q_1, Q_2; Quartili: Q_1, Q_2 (mediana), Q_3; Decili: Q_1, Q_2, …Q_9	**Differenza interquartile** Q $(Q_3 - Q_1)$	**Ordinale** **Gamma di Goodman** γ (P - Q) / (P + Q) P = nr. coppie concordanti (cioè per le quali x ed y in un caso sono entrambi maggiori o minori di un altro caso); Q = nr. coppie discordanti (cioè per le quali x cresce ed y cala rispetto ad un altro caso, o viceversa) Varia da un min di -1 (perfetta relazione positiva) ad un max di +1 (perfetta relazione negativa); se invece l'indice è 0 c'è assenza di relazione	

Cardinale

Media aritmetica

M

$(x_1+x_2+\ldots+x_n)/N = {}_{i=1}\sum^{N} x_i/N$

x_i = valori osservati; N = nr. dei valori osservati

Media aritmetica ponderata

M

$(x_1*f_1+x_2*f_2+\ldots+x_n*f_n)/N = ({}_{i=1}\sum^{N} x_i*f_i)/N$

f_i = frequenza

Media ponderata con le frequenze; se le frequenze sono raggruppate in classi si prendono i valori centrali delle classi

Media quadratica

Mq

$\sqrt{[(x_1^2+x_2^2+\ldots+x_n^2)/N]} = \sqrt{({}_{i=1}\sum^{N} x_i^2/N)}$

Anche per la Mq può essere calcolata la media ponderata

Media armonica

Ma

$N/(1/x_1+1/x_2+\ldots+1/x_n) = N/{}_{i=1}\sum^{N} 1/x_i$

Anche per la Ma può essere calcolata la media ponderata

Media geometrica

Mg

$^{N}\sqrt{(x_1*x_2*\ldots*x_n)} = {}^{N}\sqrt{({}_{i=1}\prod^{N} x_i)}$

Anche per la Mg può essere calcolata la media ponderata

Formula generale delle medie

$M_{(X)}$

$^{r}\sqrt{[{}_{i=1}\sum^{N}(x_i^{r} * f_i / N)]}$

r = 1: Media aritmetica; r = 2: Media quadratica; r = -1: Media armonica; r → 0: Media geometrica

Campo di variazione

ω

$(x_{max} - x_{min})$

Scarto Semplice Medio

SSM

$({}_{i=1}\sum^{N} | x_i - M |)/N$

Devianza

Dev

${}_{i=1}\sum^{N}(x_i - M)^2$

Varianza

Var. oppure σ^2

$\sum(x_i - M)^2/N$

Rapporto tra Devianza e N=nr. dei casi

Scarto Quadratico Medio o Deviazione Standard

SQM o DS o σ

$\sqrt{[\sum(x_i - M)^2/N]} = \sqrt{\sigma^2}$

Radice quadrata della Varianza

Coefficiente di Variazione

Cv

σ/M

Dividendo (relativizzando) la DS per la media si possono fare confronti tra gli indici; in generale si può relativizzare qualsiasi indice di variabilità

Indici di concentrazione

Curva di Lorenz

Ha un'evidenza grafica sul piano cartesiano, ottenuta riportando sugli assi rispettivamente p_i e q_i

p_i = frequenza cumulata del numero di casi; q_i = frequenza cumulata dei valori osservati

Si utilizza solo per variabili relative a quantità possedute che possono trasferirsi da un soggetto all'altro (p.es. i redditi pro-capite)

Rapporto di Concentrazione di Gini

R

$[1 - {}_{i=1}\sum^{N}(p_i - p_{i-1})*(q_i + q_{i-1})] * [N/(N-1)]$ oppure ${}_{i=1}\sum^{N-1}(p_i - q_i) / {}_{i=1}\sum^{N-1} p_i = [2/(N-1)] * {}_{i=1}\sum^{N-1}(p_i - q_i)$

Se N è grande il fattore di correzione [N/(N-1)] nella 1°formula può essere omesso (→ 1)

E' il rapporto [Area di Concentrazione/Area di max Concentrazione] del grafico di Lorenz; varia da 0 (equidistribuzione) a 1 (max concentrazione)

Indici di distanza (o di dissimilarità) oppure Misure di disuguaglianza

(misurano la distanza non tra valori e media, ma tra i valori)

Distanza Euclidea

D_{ij}

$\sqrt{[(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \ldots + (x_{in} - x_{jn})^2]}$

E' la distanza delle coppie di punti x_{ij} in uno spazio a n dimensioni

Differenza media con ripetizione

Δ_R

$({}_{ij}\sum | x_i - x_j |)/N^2$

Considera anche la differenza di un valore con se stesso (cioè i casi i=j)

Differenza media senza ripetizione

Δ

$({}_{ij}\sum | x_i - x_j |)/[N*(N-1)]$

Non considera i casi i=j

Indici di forma

(asimmetria = assenza di specularità; curtosi = altezza di una curva al suo max e lunghezza delle code)

Condizioni di Asimmetria

Simmetria se M=Mediana=Moda

Asimmetria positiva se Mo<Me<M

Asimmetria negativa se M<Me<Mo

Indice di Asimmetria di Fisher

α_F

$(1/N)*{}_{i=1}\sum^{N}[(x_i - M)/\sigma]^3$

Figure di Curtosi

Curva normale = Mesocurtica

Curva con altezza più bassa e code più lunghe = Iponormale o Platicurtica

Curva con altezza più alta e code più corte = Ipernormale o Leptocurtica

Indice di Curtosi di Pearson

β_2

$(1/N)*{}_{i=1}\sum^{N}[(x_i - M)/\sigma]^4$

Se l'indice è maggiore di 3 la distribuzione è ipernormale; se è uguale a 3 è normale; se è inferiore a 3 è iponormale

Indice di Curtosi di Fisher

γ_2

$(\beta_2 - 3)$ = (Indice di Curtosi di Pearson: β_2) - 3

Se l'indice è positivo la distribuzione è ipernormale; se è uguale a 0 è normale; se è negativo è iponormale

Diagramma di Dispersione

Possono verificarsi 4 casi:

1) relazione lineare positiva

2) relazione lineare negativa

3) nessuna relazione

4) relazione non lineare

Coefficiente di Regressione

b

Covarianza di x e y: $[\sum(x_i - M_x) * \sum(y_i - M_y)]$ / Varianza di x: $\sum(x_i - M_x)^2$

Retta di regressione (o interpolante) con il metodo dei minimi quadrati: y = a + b*x, dove b = Coeff. di regressione ed $a = M_y - b*M_x$

Se si invertono le variabili x ed y si ottengono valori diversi di a e b

Coefficiente di Correlazione di Pearson (forza della Correlazione)

r

Covarianza di x e y: $[\sum(x_i - M_x) * \sum(y_i - M_y)]$ / prodotto Dev. Standard di x e y: $\sqrt{[\sum(x_i - M_x)^2 * \sum(y_i - M_y)^2]}$

Rapporto [Covarianza / (Dev. Standard di x * Dev. Standard di y)]

Varia da un min di -1 (relaz. negativa) ad un max di +1 (relaz. positiva), con 0 = nessuna relaz.; si considera rilevante r > 0,30 (r^2 > 10%)

Si può costruire anche una "Matrice di Correlazione", ovvero la matrice dei coefficienti di correlazione di tutte le coppie di variabili

Coefficiente di Determinazione

r^2

E' il quadrato del Coeff. di Correlazione r

Rapporto (Dev. spiegata / Dev. totale); in particolare $r^2 = b_{xy} * b_{yx}$, con b = Coeff. di Regressione

Non misura la forza della correlazione, ma indica la parte di variazione della variab. dipendente y determinata dalla variazione della variab. indipendente x

Analisi inferenziale (o campionaria)

Variabili discrete

Distribuzione binomiale (o di Bernoulli)

$\{(^{n}{}_{x}) = [n! / (x!(n - x)!)]\} * p^{x} q^{n-x}$

x = variabile casuale; n = nr. prove; p = probabilità di successo; q = prob. contraria

Ipotizza che ad ogni prova la probabilità rimanga costante (la pallina è rimessa nell'urna: estrazioni con ripetizione); è lo sviluppo del binomio di Newton $(p + q)^{n}$

Media = np ; Varianza = npq

Combinazione tipica (o più probabile) della distr. binomiale

Valore: (np - q) <= x <= (np + p)

Probabilità: $[n! / (np!(n - np)!)] * p^{np} q^{n-np} = [n! / (np! * nq!)] * p^{np} q^{nq}$

Distribuzione di Poisson

$e^{-m}(m^{x} / x!)$

m = np = nr. successi

Si ottiene dalla binomiale facendo tendere p a zero ed n ad infinito

Media = m ; Varianza = m

Distribuzione ipergeometrica

$[(^{Np}{}_{x}) * (^{Nq}{}_{n-x})] / (^{N}{}_{n})$

N = nr. elementi; n = nr. elementi presi (ad n ad n)

Ipotizza che ad ogni prova la probabilità non rimanga costante (la pallina non è rimessa nell'urna: estrazioni senza ripetizione o in blocco); per N tendente ad infinito diventa la binomiale

Media = np ; Varianza = npq*[(N-n)/(N-1)]

Variabili continue

Distribuzione normale (o gaussiana, o degli errori accidentali, o "a campana")

$(1 / [\sigma\sqrt{(2\pi)}]) * \exp -[(x-m)^{2}/2\sigma^{2}]$

m = media; σ^{2} = varianza

Deriva dalla binomiale con p costante ed n che tende ad infinito

Media = m ; Varianza = σ^{2}

Distribuzione normale standardizzata

$(1 / [\sqrt{(2\pi)}]) * \exp -(z^{2}/2)$

La variabile casuale si standardizza così: z = (x-m)/σ

E' tabulata, così da conoscerne le probabilità in un certo intervallo

Media = 0 ; Varianza = 1

Distribuzione t di Student

$Y / [(1 + x^{2}/\nu)^{½(\nu+1)}]$

Y = una funzione di ν ; ν = nr. gradi di libertà

E' tabulata, così da conoscerne le probabilità in un certo intervallo

Coincide con la distr. normale quando ν = ∞

Media = 0 per ν>1 ; Varianza = ν / (ν-2) per ν>2

Distribuzione Chi-quadrato χ^{2}

$_{i=1}\Sigma^{\nu} Y^{2}$

ν = nr. gradi di libertà; Y = variabili casuali normali standardizzate e indipendenti

E' generata dalla somma di un nr. ν di v.c. normali standardizzate e indipendenti al quadrato

Al crescere di n tende alla distrib. normale

E' tabulata, così da conoscerne le probabilità in un certo intervallo

Media = ν ; Varianza = 2ν

Distribuzione F di Fisher-Snedecor

$(Y_{1}/\nu_{1}) / (Y_{2}/\nu_{2})$

$Y_{1} = \chi_{1}^{2}$; $Y_{2} = \chi_{2}^{2}$; ν_{1} = nr. gradi libertà di χ_{1}^{2} ; ν_{2} = nr. gradi di libertà di χ_{2}^{2}

E' il rapporto di 2 v.c. chi-quadrato indipendenti tra loro e divise per i rispettivi gradi di libertà

E' tabulata, così da conoscerne le probabilità in un certo intervallo

Media = $\nu_{2}/(\nu_{2}-2)$ per $\nu_{2}>2$; Varianza = $[2(\nu_{2})^{2}*(\nu_{1}+\nu_{2}-2)] / [\nu_{1}(\nu_{1}-2)^{2}*(\nu_{2}-4)]$ per $\nu_{2}>4$

Stima puntuale ed intervellare delle ipotesi

Distribuzione medie campionarie

Media = uguale alla media della popolazione

Varianza campionamento bernoulliano = σ^{2}/n

Varianza campionamento in blocco = $(\sigma^{2}/n)*[(N-n)/(N-1)]$

σ^{2} = varianza della popolazione; n = nr. elementi campione; N = nr. elementi popolazione

Test sulla media

Se la varianza del campione è nota, la statistica da utilizzare è: (x-m) / (σ/√n)

Che si distribuice, se ipotesi nulla vera, come una v.c. normale standardizzata

Se la varianza del campione non è nota, la statistica da utilizzare è: (x-m) / (S/√n)

S = varianza del campione (quale stima di quella della popolazione)

Che si distribuisce, se ipotesi nulla vera, come una v.c. t di Student con ν = n-1 gradi di libertà

Intervallo di confidenza

Se la varianza del campione è nota, l'intervallo ha i seguenti limiti: $[m - z_{a}*(\sigma/\sqrt{n})]$ e $[m + z_{a}*(\sigma/\sqrt{n})]$

Dove m = media e z_{a} = valore della v.c.normale standard. tabulato per il livello di significatività a

Se la varianza del campione non è nota, l'intervallo ha i seguenti limiti: $[m - t_{a}*(S/\sqrt{n})]$ e $[m + t_{a}*(S/\sqrt{n})]$

Dove S = varianza stimata dal campione e t_{a} valore della v.c. t di Student tabulato per il livello di significatività a

Terorema del limite centrale

Dato un universo qualsiasi con media = m e varianza = σ^{2}, la distribuzione delle medie campionarie è normale al crescere della numerosità n del campione

Pertanto la distribuzione delle medie campionarie ha media = m e varianza = σ^{2}/n

Quindi la statistica (x-m) / (σ/√n) si distribuisce come una v.c. normale standardizzata per n tendente all'infinito

In pratica si usa la tabulazione della v.c. normale standard. quando n>30, stimando la varianza dell'universo σ da quella del campione S

Numerosità del campione

In funzione dell'intervallo per le medie: $n = (4z_{a}^{2}\sigma^{2}) / I^{2}$

con z_{a} = valore tabulato per il livello a ; I = intervallo di confidenza

In funzione dell'errore di campionamento per le medie: $n = (z_{a}^{2}\sigma^{2}) / e^{2}$

con z_{a} = valore tabulato per il livello a ; e = errore di campionamento

In funzione dell'errore di campionamento per le proporzioni: $n = (z_{a}^{2}pq) / e^{2}$

con p = proporzione a favore; q = prop. contraria; la max variabilità si ha con p=q=0,50 e pertanto, dato il valore tabulato (z_{a}), si può calcolare (n) solo in funzione dell'errore (e) o dell'intervallo (I), tenendo presente che (I=2*e) e che la varianza è (pq), il cui valore max è 0,25

Per maggior precisione, quando n>5% di N, è da considerare il fattore di correzione per popolazione finita: (N-n) / (N-1)

Errore di campionamento

Errore per le medie: $e = z_{a}*\sigma/\sqrt{n}*\sqrt{(1-n/N)}$

Errore per le proporzioni: $e = z_{a}*\sqrt{(pq/n-1)}*\sqrt{(1-n/N)}$

(1-n/N) = fattore di correzione per popolazione finita, da usare quando n>0,05*N

Teoria della probabilità

Definizione classica di probabilità

P(E) = m/n

P=probabilità; E=evento; m=nr. dei possibili risultati che danno luogo all'evento (E); n=nr. di tutti i possibili risultati

0<=P(E)<=1

dove: 0=(E) impossibile e 1=(E) certo

Definizione frequentista di probabilità (legge dei grandi numeri)

$P(E) = \lim_{n\to\infty}[fr_{n}(E)/n]$

$fr_{n}(E)$=nr. di volte che in un esperimento si è verificato (E); n=nr. degli esperimenti effettuati

Definizione soggettivista di probabilità

P(E) = somma che un individuo razionale sarebbe disposto a scommettere in un gioco equo per ricevere un'unità di vincita nel caso si verifichi (E)

Probabilità condizionata

P(A/E) = P(AE)/P(E)

probabilità che si verifichi (A) dato il verificarsi di (E)

Teorema delle probabilità totali

P(A+B+C+… …+Z) = P(A)+P(B)+P(C)+… …+P(Z)

per un nr. n di eventi (A, B, C, …) che si escludono a vicenda, aventi prob. P(A), P(B), P(C)… di verificarsi, la prob. che si verifichi uno qualunque di essi è data dalla somma delle rispettive probabilità

Teorema delle probabilità composte

P(A*B*C*… …*Z) = P(A)*P(B)*P(C)*… …*P(Z)

per un nr. n di eventi (A, B, C, …) indipendenti tra di loro, aventi prob. P(A), P(B), P(C)… di verificarsi, la prob. che tutti si verifichino è data dal prodotto delle rispettive probabilità

Teorema di Bayes

$P(A_{i}/E) = [P(E/A_{i})*P(A_{i})] / P(E) = [P(E/A_{i})*P(A_{i})] / [_{i=1}\Sigma^{n} P(E/A_{i})*P(A_{i})]$

Arricchisce la quantificazione delle probabilità dell'esperimento (A) con le n osservazioni dell'evento (E)

Calcolo combinatorio

Disposizioni con ripetizioni

$D_{n,m} = n^{m}$

Il nr. di gruppi di m oggetti ordinati, anche ripetuti, che si possono formare con un insieme di n elementi

Disposizioni senza ripetizioni

$D_{n,m} = n!/(n-m)! = n*(n-1)*(n-2)*… …*(n-m+1) = n! / (n-m)!$

Il nr. di gruppi di m oggetti ordinati, non ripetuti, che si possono formare con un insieme di n elementi

Permutazioni senza ripetizioni

$P_{n} = n!$

Disposizioni senza ripetizioni di n oggetti presi da n elementi (cioè $D_{n,m}$ con n=m)

Permutazioni con ripetizioni (con gruppi di elementi che si ripetono)

$P_{m} = m!/(m_{1}!*m_{2}!*… …*m_{n}!)$

I gruppi sono formati rispettivamente da m_{1}, m_{2}, …m_{n} elementi tutti uguali fra di loro

Combinazioni con ripetizioni

$C_{n,m} = C_{n+m-1,m} = (^{n+m-1}{}_{m})$

Il nr. di gruppi (per i quali non ha importanza l'ordine) di m oggetti, anche ripetuti, che si possono formare con un insieme di n elementi

Combinazioni senza ripetizioni

$C_{n,m} = D_{n,m}/P_{m} = (^{n}{}_{m}) = (^{n}{}_{n-m}) = [n*(n-1)*(n-2)*… …*(n-m+1)] / m! = n! / [m!*(n-m)!]$

Il nr. di gruppi (per i quali non ha importanza l'ordine) di m oggetti, non ripetuti, che si possono formare con un insieme di n elementi