



UNIVERSITÀ DI PERUGIA  
Dipartimento di Matematica e Informatica



Appunti *Simulazione*

# Formulario

---

Anno Accademico 2021-2022

*Last Update: January 18, 2023*

# Contents

<b>1</b>	<b>Distribuzioni</b>	<b>4</b>
1.1	Stimare la Distribuzione . . . . .	4
1.2	Calcolare la Probabilità di una Distribuzione . . . . .	5
1.2.1	Esponenziale . . . . .	5
1.2.1.1	Senza Intervalli . . . . .	5
1.2.1.2	Con Intervalli . . . . .	5
1.2.2	Poisson . . . . .	5
1.2.3	Geometrica . . . . .	5
<b>2</b>	<b>Goodness of Fit</b>	<b>6</b>
2.1	Test $\chi^2$ . . . . .	6
2.1.1	Dati senza Intervalli . . . . .	6
2.1.2	Dati con Intervalli . . . . .	7
2.2	Test Kolmogorov . . . . .	8
2.2.1	Dati Senza Intervalli . . . . .	9
2.2.2	Dati Con Intervalli . . . . .	9
2.3	Informazioni utili su Formule . . . . .	10
2.3.1	Komorov . . . . .	10
2.3.1.1	<i>cumsum</i> . . . . .	10
2.4	Tabelle di Riferimento . . . . .	11
2.4.1	Tabella di Riferimento Test $\chi^2$ . . . . .	11
2.4.2	Tabella di Riferimento Test Kolmogorov . . . . .	12



*”Oi, con quanto sentimento  
defeco sul tuo naso,  
così che ti coli sul mento.”*

Wolfgang Amadeus Mozart

# Chapter 1

## Distribuzioni

### 1.1 Stimare la Distribuzione

Per stimare una distribuzione avendo solo i dati iniziali del problema effettua le seguenti operazioni:

**N.B.** nel caso di *Intervalli*,  $categoria_i$  va sostituito con *Punto Medio Intervallo<sub>i</sub>*

1.  $n = \sum f_i$ : assicurati di aver calcolato la somma totale delle osservazioni

2. Calcola la **Media**:

(a) Aggiungi *Colonna Totale*:  $categoria_i * f_i$

(b) Calcola la media effettiva con:  $media = \frac{\sum totale}{n}$

3. Calcola la **Varianza**  $\sigma^2$ :

(a) Aggiungi *Colonna ris*:  $(categoria_i - media)^2 * f_i$

(b) Calcola la varianza effettiva con:  $\sigma^2 = \frac{\sum ris}{n-1}$

4. Calcola la **Deviazione Standard**  $\sigma$ :

(a)  $\sigma = \sqrt{\sigma^2}$

5. Calcola  $V = \frac{\sigma}{media}$

Una volta completati tutti i calcoli controlla se il coefficiente  $V$  è vicino ad 1 e:

- **se lo è** allora utilizza l'**Esponenziale**,
- **se non lo è** è **Poissoniana** ma per avere una verifica, controlla che la media e la varianza siano uguali.



**Note:**

Si può avere una prima idea del tipo di distribuzione anche osservando le frequenze per categoria:

- Se le frequenze hanno valori alti per le prime categorie e poi decrescono, probabilmente è esponenziale negativa.
- Se le frequenze hanno valori bassi per le prime categorie e poi crescono, probabilmente è esponenziale...
- Se le frequenze hanno valori alti nelle categorie centrali e bassi verso le categorie agli estremi, probabilmente è Poissoniana
- Nel caso in cui sia geometrica solitamente viene esplicitato.

## 1.2 Calcolare la Probabilità di una Distribuzione

### 1.2.1 Esponenziale

#### 1.2.1.1 Senza Intervalli

$$p(i) = \frac{e^{\frac{-\text{categoria}_i}{\text{media}}}}{\text{media}}$$

#### 1.2.1.2 Con Intervalli

$$p(i) = 1 - e^{\frac{-\text{categoria}_i}{\text{media}}}$$

### 1.2.2 Poisson

$$p(i) = \frac{e^{-\text{media}} * \text{media}^{\text{categoria}_i}}{\text{categoria}_i!}$$

### 1.2.3 Geometrica

$$p(i) = \rho * (1 - \rho)^{\text{categoria}_i}$$



# Chapter 2

## Goodness of Fit

### 2.1 Test $\chi^2$

Devi utilizzare questa sezione solo se il numero delle **osservazioni totale**  $n > 30$ .

#### 2.1.1 Dati senza Intervalli

Devi utilizzare questa sezione solo quando hai dei dati **Senza Intervalli**, devi anche fare attenzione che il **numero di osservazioni**  $n > 30!!$

Operazioni da effettuare:

1. Riportare i dati in una tabella in Calc:
  - *Colonna 1: categorie*
  - *Colonna 2:  $f_i$*
2. Raggruppare le categorie se  $\exists \text{ categoria} < 5$ :
  - Parti dall'ultimo a salire (dal basso verso l'alto delle categorie)
  - Raggruppare tutte nell'ultima categoria che le faccia diventare maggiori di 5 sommando le frequenze.
  - *Esempio:*



	A	B	C	D	E
1	VALORI	frequenze	f(i) raggruppate		
2	0	59	59		
3	1	26	26		
4	2	24	24		
5	3	18	18		
6	4	12	12		
7	5	5	5		
8	6	4	12		
9	7	3			
10	9	3			
11	11	2			
12					

	A	B	C	D
1	VALORI	frequenze	f(i) raggruppate	
2	0	59	59	
3	1	26	26	
4	2	24	24	
5	3	18	18	
6	4	12	12	
7	5	5	9	
8	6	1		
9	7	1		
10	9	1		
11	11	1		

3. Calcolare:

- $n = \sum(f_i)$
- $f(i) = f_i/n$ : non serve
- Capire la distribuzione se non è data (vedi 1.1)
- $p(i)$ : dipende dalla distribuzione (vedi 1.2)
- $F_i = n * p(i)$ : numero di intervalli unitari teorici con  $i$  arrivi
- $G_i = \frac{(f_i - F_i)^2}{F_i}$
- $V = \sum G_i$ : sommare tutti i valori di  $G$
- $df = \text{Numero Categorie} - 1 - \text{Numero Parametri Distribuzione}$

Una volta terminati i calcoli devi guardare la riga nella tabella del  $\chi^2$  (AGGIUNGERE REF) con lo stesso valore di  $df$ : devi controllare che il valore  $V$  ricada negli intervalli che non superino il  $P_{95}$ .

## 2.1.2 Dati con Intervalli

Devi utilizzare questa sezione solo quando hai dei dati divisi in **Intervalli**, devi anche fare attenzione che il **numero di osservazioni**  $n > 30$ !!

Calcoli da effettuare:

- Riportare i dati in una tabella in Calc:
  - Colonna 1: categorie*, probabilmente devi aggiungerle tu, parti da 0 in poi
  - Colonna 2: intervallo*, del tipo  $x_1 - x_2$ . Fai sempre attenzione che  $x_2 \geq x_1$  !!! In caso li inverti.
  - Colonna 3: frequenza*  $f_i$



2. Aggiungere *Colonna*  $x_1$  (intervallo più piccolo)
3. Aggiungere *Colonna*  $x_2$  (intervallo più grande)
4. Aggiungere *Colonna Punto Medio Intervalli* tra  $x_2$  e  $x_1$  con  $\frac{x_1+x_2}{2}$
5. Calcolare:
  - (a) capire la distribuzione se non è data (vedi 1.1)
  - (b)  ~~$f(i) = f_i/n$ : non serve~~
  - (c)  $p(i) = p(x_2) - p(x_1)$  = calcolare secondo la distribuzione (vedi 1.2)
  - (d)  $F_i = n * p(i)$ : numero di intervalli unitari teorici con  $i$  arrivi
  - (e)  $G_i = \frac{(f_i - F_i)^2}{F_i}$
  - (f)  $V = \sum G_i$ : sommare tutti i valori di  $G$
  - (g)  $df$  = Numero Categorie  $- 1 -$  Numero Parametri Distribuzione
6. Raggruppare le categorie se  $\exists$  categoria  $< 5$ :
  - Parti dall'ultimo a salire (dal basso verso l'alto delle categorie)
  - Raggruppare tutte nell'ultima categoria che le faccia diventare maggiori di 5 sommando le frequenze.
  - *Esempio:*

	A	B	C	D	E
1	VALORI	frequenze	f(i) raggruppate		
2	0	59	59		
3	1	26	26		
4	2	24	24		
5	3	18	18		
6	4	12	12		
7	5	5	5		
8	6	4	12		
9	7	3			
10	9	3			
11	11	2			
12					

	A	B	C	D
1	VALORI	frequenze	f(i) raggruppate	
2	0	59	59	
3	1	26	26	
4	2	24	24	
5	3	18	18	
6	4	12	12	
7	5	5	9	
8	6	1		
9	7	1		
10	9	1		
11	11	1		

Una volta terminati i calcoli devi guardare la riga nella tabella del  $\chi^2$  (AGGIUNGERE REF) con lo stesso valore di  $df$ : devi controllare che il valore  $V$  ricada negli intervalli che non superino il  $P_{95}$ .

## 2.2 Test Kolmogorov

Devi utilizzare questa sezione solo se il numero delle **osservazioni totale**  $n < 30$ .





### 2.2.1 Dati Senza Intervalli

Devi utilizzare questa sezione solo quando hai dei dati **Senza Intervalli**, devi anche fare attenzione che il **numero di osservazioni totali**  $n < 30!!$

Operazioni da effettuare:

1. Riportare i dati in una tabella in Calc:

- *Colonna categorie*
- *Colonna frequenze  $f_i$*

2. Calcolare:

- (a)  $f(i) = f_i/n$ : frequenze osservate
- (b) Individuare la distribuzione di probabilità adatta (vedi 1.1)
- (c)  $p(i)$ : probabilità teorica (vedi 1.2)
- (d)  $d_i = \text{cumsum}(f(i))$ : somma cumulativa delle  $f(i)$
- (e)  $D_i = \text{cumsum}(p(i))$ : somma cumulativa delle  $p(i)$
- (f)  $D = |d_i - D_i|$ : la differenza assoluta
- (g)  $D_{max} = \max(D)$ : il massimo valore tra le differenze assolute  $D$

Una volta completati tutti i calcoli, cercare nella tabella di *Kolmogorov-Smirnov* la riga corrispondente al valore delle osservazioni totali  $n$ : se il valore  $D_{max}$  è sotto il  $D_{0,10}$  la distribuzione è accettata, altrimenti no.

### 2.2.2 Dati Con Intervalli

Devi utilizzare questa sezione solo quando hai dei dati **Senza Intervalli**, devi anche fare attenzione che il **numero di osservazioni totali**  $n < 30!!$

**N.B.:** *non abbiamo trovato esercizi con cui testare questa sezione !*

Operazioni da effettuare:

1. Riportare i dati in una tabella in Calc:

- *Colonna categorie*: probabilmente devi aggiungerle tu, parti da 0 in poi
- *Colonna intervallo*: del tipo  $x_1 - x_2$ . Fai sempre attenzione che  $x_2 \geq x_1$  !!! In caso li inverti.



- *Colonna frequenze  $f_i$*
- 2. Aggiungere *Colonna  $x_1$*  (estremo più piccolo dell'intervallo)
- 3. Aggiungere *Colonna  $x_2$*  (estremo più grande dell'intervallo)
- 4. Calcolare:
  - (a)  $f(i) = f_i/n$ : frequenze osservate
  - (b) Individuare la distribuzione di probabilità adatta (vedi 1.1)
  - (c)  $p(i) = p(x_2) - p(x_1)$ : probabilità teorica per ogni intervallo (vedi 1.2)
  - (d)  $d_i = \text{cumsum}(f(i))$ : somma cumulativa delle  $f(i)$
  - (e)  $D_i = \text{cumsum}(p(i))$ : somma cumulativa delle  $p(i)$
  - (f)  $D = |d_i - D_i|$ : la differenza assoluta
  - (g)  $D_{max} = \max(D)$ : il massimo valore tra le differenze assolute  $D$

Una volta completati tutti i calcoli, cercare nella tabella di *Kolmogorov-Smirnov* la riga corrispondente al valore delle osservazioni totali  $n$ : se il valore  $D_{max}$  è sotto il  $D_{0,10}$  la distribuzione è accettata, altrimenti no.

## 2.3 Informazioni utili su Formule

### 2.3.1 Komorov

#### 2.3.1.1 *cumsum*

Per calcolare *cumsum* (somma cumulativa) va eseguito il seguente procedimento:

- La prima cella resta uguale alla prima cella della colonna di riferimento (es.  $f(i)$  o  $p(i)$ )
- Dalla seconda cella in poi si blocca la prima cella della somma cumulativa (quella calcolata al punto precedente) e si somma fino alla cella  $i$  di riferimento (vedi Figura 2.1)

$f_k$	=C2						
	B	C	D	E	F	G	
	FREQUENZE	f(i)	totale	ris	p(i)	d_i	
0	3	0,15	0	10,83	0,149569	0,15	
1	6	0,3	6	4,86	0,28418	0,45	
2	5	0,25	10	0,05	0,269971	0,7	
3	3	0,15	9	3,63	0,170982	0,85	
4	2	0,1	8	8,82	0,081216	0,95	
5	1	0,05	5	9,61	0,030862	1	

  

	B	C	D	E	F	G	
	FREQUENZE	f(i)	totale	ris	p(i)	d_i	
	3	0,15	0	10,83	0,149569	0,15	
	6	0,3	6	4,86	0,28418	0,45	
	5	0,25	10	0,05	0,269971	0,7	
	3	0,15	9	3,63	0,170982	0,85	
	2	0,1	8	8,82	0,081216	0,95	
	1	0,05	5	9,61	0,030862	1	

  

	B	C	D	E	F	G	
	FREQUENZE	f(i)	totale	ris	p(i)	d_i	
	3	0,15	0	10,83	0,149569	0,15	
	6	0,3	6	4,86	0,28418	0,45	
	5	0,25	10	0,05	0,269971	0,7	
	3	0,15	9	3,63	0,170982	0,85	
	2	0,1	8	8,82	0,081216	0,95	
	1	0,05	5	9,61	0,030862	1	

Figure 2.1: Esempio di calcolo della funzione *cumsum*

## 2.4 Tabelle di Riferimento

### 2.4.1 Tabella di Riferimento Test $\chi^2$

Tabella 2.9 Percentili della distribuzione $\chi^2$										
df	$P_{0,5}$	$P_1$	$P_{2,5}$	$P_5$	$P_{10}$	$P_{90}$	$P_{95}$	$P_{97,5}$	$P_{99}$	$P_{99,5}$
1	0,000039	0,00016	0,00098	0,0039	0,0158	2,71	3,84	5,02	6,63	7,88
2	0,0100	0,0201	0,0506	0,1026	0,2107	4,61	5,99	7,38	9,21	10,60
3	0,0717	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34	12,84
4	0,207	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,412	0,554	0,831	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,676	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,989	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
24	9,89	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
30	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67
40	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77
60	35,53	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38	91,95
120	83,85	86,92	91,58	95,70	100,62	140,23	146,57	152,21	158,95	163,64

Figure 2.2: Tabella di Riferimento per Test  $\chi^2$



### 2.4.2 Tabella di Riferimento Test Kolmogorov

$n$	$D_{0,10}$	$D_{0,05}$	$D_{0,01}$
1	0,950	0,975	0,995
2	0,776	0,842	0,929
3	0,642	0,708	0,828
4	0,564	0,624	0,733
5	0,510	0,565	0,669
6	0,470	0,521	0,618
7	0,438	0,486	0,577
8	0,411	0,457	0,543
9	0,388	0,432	0,514
10	0,368	0,410	0,490
11	0,352	0,391	0,468
12	0,338	0,375	0,450
13	0,325	0,361	0,433
14	0,314	0,349	0,418
15	0,304	0,338	0,404
16	0,295	0,328	0,392
17	0,286	0,318	0,381
18	0,278	0,309	0,371
19	0,272	0,301	0,363
20	0,264	0,294	0,356
25	0,24	0,27	0,32
30	0,22	0,24	0,29
35	0,21	0,23	0,27
Oltre 35	1,22 $\sqrt{n}$	1,36 $\sqrt{n}$	1,63 $\sqrt{n}$

Figure 2.3: Tabella di Riferimento per Test Kolmogorov