



UNIVERSITA' DEGLI STUDI DI VERONA

LABORATORIO DI PROBABILITA' E STATISTICA

Docente: Bruno Gobbi

5 - ESERCIZI RIEPILOGATIVI PRIME 3 LEZIONI

ES. LAUREATI PER FACOLTA' E LAVORO

ESERCIZIO 5: La tabella presenta la distribuzione di alcuni laureati in diverse facoltà e la loro condizione occupazionale dopo un anno dalla laurea.

CONDIZIONE OCCUPAZIONALE	FACOLTA'		
	Informatica	Lettere	Giurisprudenza
Occupati	70	40	30
In cerca di occupazione	20	15	30

g.d.l.	alpha (significatività)	
	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. LAUREATI PER FACOLTA' E LAVORO

```
> laureati=matrix(c(70, 40, 30, 20, 15, 30), nrow=2,  
byrow=TRUE)  
> condizione=c("Occupati", "In cerca di occupazione")  
> facoltà=c("Informatica", "Lettere", "Giurisprudenza")  
> dimnames(laureati)=list(condizione, facoltà)  
> laureati
```

	Informatica	Lettere	Giurisprudenza
Occupati	70	40	30
In cerca di occupazione	20	15	30

```
> mosaicplot(laureati)
```

ES. LAUREATI PER FACOLTA' E LAVORO

```
> testchiq=chisq.test(laureati)
> testchiq
```

Pearson's Chi-squared test

data: laureati

X-squared = 13.5108, df = 2, p-value = 0.001165

POICHE' IL VALORE CALCOLATO DEL CHI-QUADRATO E' 13.5108, SUPERIORE ALLA SOGLIA CRITICA DI 9,21 VALIDO ALL'1% PER 2 G.D.L., SI RIFIUTA L'IPOTESI NULLA DI INDIPENDENZA E SI CONFERMA LA CONNESSIONE FRA I FENOMENI

I GRADI DI LIBERTA' SONO 2 PERCHE' DATI DA $(r-1)*(c*1)=(2-1)*(3-1)$

g.d.l.	alpha (significatività)	
	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. LAUREATI PER FACOLTA' E LAVORO

CALCOLIAMO IL VALORE DELLA STATISTICA V DI CRAMER

```
> chiquadrato= testchiq$statistic  
> chiquadrato  
X-squared  
13.51079
```

IL TOTALE DI ELEMENTI PRESENTI SI OTTIENE IN QUESTO MODO:

```
> N = sum(laureati)  
> N  
[1] 205
```

SI SCEGLIE IL MINORE FRA IL NUMERO DI RIGHE E DI COLONNE E SI SOTTRAE 1

```
> V=sqrt( chiquadrato / (N*(2-1)) )  
> V  
X-squared  
0.2567223
```

IL RISULTATO PORTA AD AFFERMARE CHE C'È UNA DISCRETA CONNESSIONE FRA I DUE FENOMENI

ES. STAGE E ASSUNZIONE (CASO NORMALE)

ESERCIZIO 7 A: Si vuole verificare se esiste una relazione fra il fatto di svolgere uno stage presso un importante istituto di credito e la successiva eventuale assunzione. Sono stati così presi in considerazione 200 ragazzi così distribuiti:

		ASSUNZIONE?		
		SI'	NO	Totale
STAGE?	SI'	80	20	100
	NO	25	75	100
	Totale	105	95	200

g.d.l.	alpha (significatività)	
	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. STAGE E ASSUNZIONE (CASO NORMALE)

```
> stage_lavoro=matrix(c(80, 20, 25, 75), nrow=2,  
byrow=TRUE)
```

```
> stage=c("sì stage", "no stage")
```

```
> lavoro=c("Sì assunzione", "No assunzione")
```

```
> dimnames(stage_lavoro)=list(stage, lavoro)
```

```
> stage_lavoro
```

	Sì assunzione	No assunzione
sì stage	80	20
no stage	25	75

```
> mosaicplot(stage_lavoro)
```

ES. STAGE E ASSUNZIONE (CASO NORMALE)

```
> testchiq=chisq.test(stage_lavoro)
> testchiq
```

Pearson's Chi-squared test with Yates' continuity correction

```
data: stage_lavoro
```

```
X-squared = 58.4662, df = 1, p-value = 2.068e-14
```

POICHE' IL VALORE CALCOLATO DEL CHI-QUADRATO E' 58.4662, BEN SUPERIORE ALLA SOGLIA CRITICA DI 6.64 VALIDO ALL' 1%, SI RIFIUTA L'IPOTESI NULLA DI INDIPENDENZA E SI CONFERMA LA CONNESSIONE FRA I FENOMENI, OVVERO FARE UNO STAGE COMPORTA MAGGIORI PROBABILITA' DI ESSERE ASSUNTI. I GRADI DI LIBERTA' SONO 1 PERCHE' DATI DA $(r-1)*(c*1)=(2-1)*(2-1)$

g.d.l.	alpha (significatività)	
	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. STAGE E ASSUNZIONE (CASO NORMALE)

CALCOLIAMO IL VALORE DELLA STATISTICA V DI CRAMER

```
> chiquadrato=testchisq$statistic  
> chiquadrato  
X-squared  
58.46617
```

IL TOTALE DI ELEMENTI PRESENTI SI OTTIENE IN QUESTO MODO:

```
> N = sum(stage_lavoro)  
> N  
[1] 200
```

SI SCEGLIE IL MINORE FRA IL NUMERO DI RIGHE E DI COLONNE E SI SOTTRAE 1

```
> V=sqrt( chiquadrato / (N*(2-1)) )  
> V  
X-squared  
0.5406763
```

IL RISULTATO PORTA AD AFFERMARE CHE C'È UNA BUONA CONNESSIONE FRA I DUE FENOMENI

ES. STAGE E ASSUNZIONE (CASO LIMITE 1)

ESERCIZIO 7 B: Si vuole verificare se esiste una relazione fra il fatto di svolgere uno stage presso un importante istituto di credito e la successiva eventuale assunzione. Sono stati così presi in considerazione 200 ragazzi così distribuiti:

		ASSUNZIONE?		
		SI'	NO	Totale
STAGE?	SI'	100	0	100
	NO	0	100	100
	Totale	100	100	200

g.d.l.	alpha (significatività)	
	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. STAGE E ASSUNZIONE (CASO LIMITE 1)

```
> stage_lavoro=matrix(c(100, 0, 0, 100), nrow=2, byrow=TRUE)
```

```
> dimnames(stage_lavoro)=list(stage, lavoro)
```

```
> testchiq=chisq.test(stage_lavoro)
```

```
> testchiq
```

```
data: stage_lavoro
```

```
X-squared = 196.02, df = 1, p-value < 2.2e-16
```

```
> chiquadrato=testchiq$statistic
```

```
> V=sqrt( chiquadrato / (N*(2-1)) )
```

```
> V
```

```
0.99
```

QUI C'E' LA MASSIMA CONNESSIONE, NEL SENSO CHE QUANDO UNO STUDENTE FA LO STAGE, VIENE SEMPRE ASSUNTO E VICEVERSA.

IL CHI-QUADRATO E' MOLTO ALTO (196.02) E DI CONSEGUENZA IL V DI CRAMER E' VICINISSIMO A 1 (0.99)

ES. STAGE E ASSUNZIONE (CASO LIMITE 2)

ESERCIZIO 7 C: Si vuole verificare se esiste una relazione fra il fatto di svolgere uno stage presso un importante istituto di credito e la successiva eventuale assunzione. Sono stati così presi in considerazione 200 ragazzi così distribuiti:

		ASSUNZIONE?		
		SI'	NO	Totale
STAGE?	SI'	50	50	100
	NO	50	50	100
	Totale	100	100	200

g.d.l.	alpha (significatività)	
	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. STAGE E ASSUNZIONE (CASO LIMITE 2)

```
> stage_lavoro=matrix(c(50, 50, 50, 50), nrow=2, byrow=TRUE)
```

```
> dimnames(stage_lavoro)=list(stage, lavoro)
```

```
> testchiq=chisq.test(stage_lavoro)
```

```
> testchiq
```

```
data: stage_lavoro
```

```
X-squared = 0, df = 1, p-value = 1
```

```
> chiquadrato=testchiq$statistic
```

```
> V=sqrt( chiquadrato / (N*(2-1)) )
```

```
> V
```

```
0
```

NEL CASO DI EQUIDISTRIBUZIONE, NON C'E' NESSUNA CONNESSIONE, NEL SENSO CHE I DUE FENOMENI NON SEMBRANO AVERE ALCUN EFFETTO L'UNO SULL'ALTRO. CHE UNO STUDENTE FACCIA O MENO LO STAGE, NON SEMBRA CAMBIARE LE SUE POSSIBILITA' DI ESSERE ASSUNTO. IL CHI-QUADRATO E' PARI A ZERO E DI CONSEGUENZA LO E' ANCHE IL V DI CRAMER.