Lösningar till övningar på χ^2-test

1. Data:

Intervall	*Obs*	*Antal obs*
(26, 50)	48	1
(51, 75)	52, 60, 61, 70, 71, 72, 74	7
(76, 100)	86, 86	2

Hypotesen är $\begin{cases} H_0 : \text{antalen fördelar sig jämnt} \\ H_1 : \text{antalen fördelar sig ojämnt} \end{cases}$ och sign.nivån 0.05.

Under H_0 är det förväntade antalet inom varje intervall i:

$E_i = n \cdot \dfrac{1}{\text{antal klasser}} = 10 \cdot \dfrac{1}{3} = 3.33$

så hypotesen kan därför formuleras

$\begin{cases} H_0 : E(O_i) = 3.33 \text{ för alla } i \\ H_1 : E(O_i) \neq 3.33 \text{ för något } i \end{cases}$

och vi har r-tabellen (där $r = 3$)

i	*Intervall*	O_i	E_i
1	(26, 50)	1	3.33
2	(51, 75)	7	3.33
3	(76, 100)	2	3.33

$$\begin{aligned} u &= \sum_{i=1}^{r} \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(1-3.33)^2}{3.33} + \frac{(7-3.33)^2}{3.33} + \frac{(2-3.33)^2}{3.33} \\ &= 6.2 \end{aligned}$$

$\chi^2_\alpha(r-1) = \chi^2_{0.05}(2) = 5.99 < 6.2$ varmed H_0 förkastas på sign.nivå 0.05 dvs <u>antalen fördelar sig ojämnt över intervallen.</u>

2.

Disk-medel	%
A	35
B	38
C	27

300 st. $\Rightarrow$

r-tabell ($r = 3$):

	O_i	E_i
A	105	100
B	114	100
C	81	100

Väljer $\alpha = 0.05$

$$
\begin{aligned}
u &= \sum_{i=A,B,C} \frac{(O_i - E_i)^2}{E_i} \\
&= \frac{(105-100)^2}{100} + \frac{(114-100)^2}{100} + \frac{(81-100)^2}{100} \\
&= \tfrac{1}{100}(25 + 196 + 361) \\
&= 5.82
\end{aligned}
$$

$\chi^2_\alpha(n-1) = \chi^2_{0.05}(2) = 5.991$

så H_0 kan ej förkastas på sign. nivå 0.05 (el. lägre)

dvs <u>man kan **ej** visa skillnad mellan vilket diskmedel som föredras.</u>

3. 178 invånare blev äldre än 90 år och 545 äldre än 80, så $545 - 178 = 367$ blev mellan 80 och 90 år. På samma sätt fås att $869 - 545 = 324$ blev mellan 70 och 80.
Vidare, under nollhypotesen att livslängderna är normalfördelade med $\mu = 80$ och $\sigma = 10$, är det förväntade antalet inom respektive åldersklass:
$n\,P(X < 70) = 1000\Phi(\frac{70-80}{10}) = 158.7$
$n\,P(70 < X < 80) = 1000\Big(\Phi(\frac{80-80}{10}) - \Phi(\frac{70-80}{10})\Big) = 341.3$
$n\,P(80 < X < 70) = \{\text{symmetri kring } \mu = 80\} = 341.3$
$n\,P(X > 90) = 158.7$
så r-tabellen blir

Klass	O_i	E_i
$x < 70$	131	158.7
$70 < x < 80$	324	341.3
$80 < x < 90$	367	341.3
$x > 90$	178	158.7

Hypotesen, för att testa om Blomstermålas livslängder skiljer sig från landets i övrigt, är

$$\begin{cases} H_0 : E(O_i) = E_i \text{ alla } i \\ H_1 : E(O_i) \neq E_i \text{ något } i \end{cases} \quad \text{och signifikansnivån } \alpha = 0.05.$$

$$\begin{aligned} u &= \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(131-158.7)^2}{158.7} + \frac{(324-341.3)^2}{341.3} + \frac{(367-341.3)^2}{341.3} + \frac{(178-158.7)^2}{158.7} \\ &= 4.85 + 0.88 + 1.93 + 2.35 \\ &= 10.00 \end{aligned}$$

$\chi^2_\alpha(n-1) = \chi^2_{0.05}(3) = 7.81$
Eftersom $u = 10 > 7.81 = \chi^2_{0.05}(3)$ förkastas H_0 på 5% sign.nivå, dvs
ja, livslängderna bland Blomstermålas befolkning skiljer sig från resten av landet.

4.

Antal bilar	≤ 5	$6-10$	$11-15$	≥ 16
O_i	21	116	77	10

$n = 21 + 116 + 77 + 10 = 224.$

Vi vill testa "antalet röda bilar är $Poi(9)$" mot att "antalet röda bilar är ej $Poi(9)$" dvs

$$\begin{cases} H_0 : F = Poi(9) \\ F \neq Poi(9) \end{cases} \quad \text{med sign.nivå } \alpha = 0.01.$$

Låt X = "antalet bilar under en timme en vinterdag mellan Frösakull och Högskolan". Under H_0 är då $E_1 = n \cdot P(X \leq 5) = 224 \cdot 0.11569 = 25.9$ (där sannolikheten $P(X \leq 5)$ fås från tabell eller genom egen uträkning: $e^{-9}(\frac{9^0}{0!} + \frac{9^1}{1!} + \frac{9^2}{2!} + \frac{9^3}{3!} + \frac{9^4}{4!} + \frac{9^5}{5!}) = 0.11569$).

På samma sätt är

$E_2 = 224(0.70599 - 0.11569) = 132.2$

$E_3 = 224(0.97796 - 0.70599) = 60.9$

$E_4 = 224(1 - 0.97796) = 4.9$

och den tredje raden i tabellen blir

E_i	25.9	132.2	60.9	4.9

varmed statistikan U:s värde kan beräknas

$$\begin{aligned} u &= \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(21-25.9)^2}{25.9} + \frac{(116-132.2)^2}{132.2} + \frac{(77-60.9)^2}{60.9} + \frac{(10-4.9)^2}{4.9} \\ &= 0.932 + 1.991 + 4.224 + 5.192 \\ &= 12.36 \end{aligned}$$

$\chi^2_{0.01}(4-1) = 11.3$ och $u = 12.36 > 11.3$ så H_0 förkastas

dvs ja, läraren har fel!