2014 Ext Ans.

1.

a) When $n=14$, LQ is median of lower 7 values,

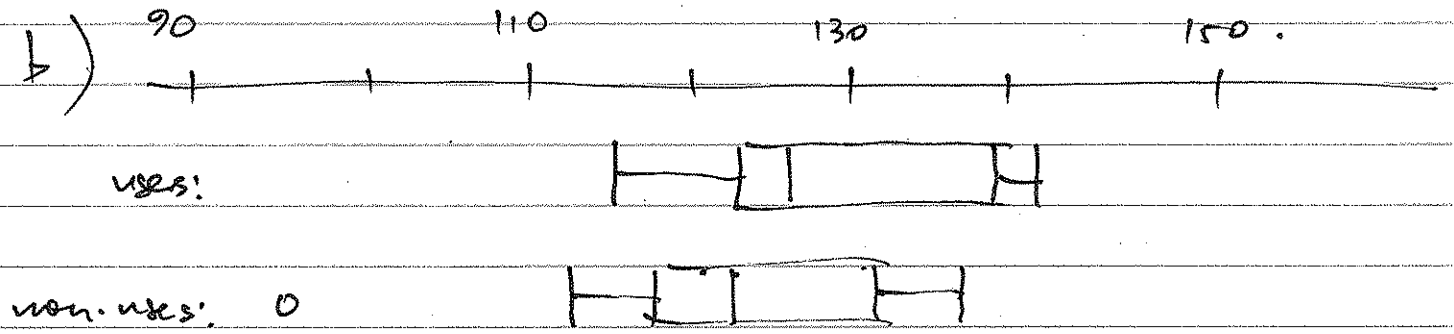
So the 4th order stat : $\underline{\underline{117.7}}$

UQ is 4th from the top: 131.0.

median is $\dfrac{x_{(7)}+x_{(8)}}{2} = \dfrac{120.1+126.5}{2} = \underline{\underline{123.3}}$

$$\boxed{94.7 \quad 117.7 \quad 123.3 \quad 131.0 \quad 138.7}$$

b)



c) Yes. The Boxplots have similar spread, both are reasonably symmetric.

d) Writing x for users, y for non-users, the C.I. takes the form.

$\bar{x}-\bar{y} \pm c \cdot s_p\sqrt{\frac{1}{10}+\frac{1}{14}}$ where c satisfies.

$P(t_{22}>c) = 0.025$.

$\Rightarrow c = 2.074$

Note $\bar{x} = 129.4$

$\bar{y} = 122.7$

$s_p = 10.29$.

this gives $6.7 \pm (2.074 \times 10.29 \times 0.414)$

i.e. 6.7 ± 8.835

or $(-2.135, 15.535)$.

2a) $r = \dfrac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \dfrac{1927}{\sqrt{1291 \times 13901}} \simeq \underline{\underline{0.454}}$

b) First note that

$$\hat{\sigma}^2 = \frac{1}{n-2} S_{yy}\left[1 - \frac{S_{xy}^2}{S_{xx} S_{yy}}\right] = \frac{S_{yy}}{n-2}\left[1 - r^2\right].$$

So $t = \dfrac{b}{se(b)} = \dfrac{S_{xy} / S_{xx}}{\hat{\sigma} / \sqrt{S_{xx}}}$

$$= \frac{S_{xy} / \sqrt{S_{xx}}}{\sqrt{\frac{S_{yy}}{n-2}(1-r^2)}}$$

$$= \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

c) Obs value of the statistic (using (b)).

$$t_{obs} \simeq \frac{0.454 \times 6}{\sqrt{1 - 0.207}} \simeq 3.059$$

So the (one-sided) p-value is $P(t_{36} \geq 3.059)$
Since ~~this~~ 3.059 is between 2.990 and 3.333 the p-value is between 0.0025 and 0.001.

3. a). i). $E(X) = \sum_x x . P(X=x)$

$= \frac{1}{6}\left[1+2+3+4+5+6\right] = \frac{21}{6} = \frac{7}{2} = 3.5$

ii) $Var(X) = E(X^2) - \left[E(X)\right]^2$

$E(X^2) = \sum_x x^2 P(X=x)$

$= \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}.$

So $Var(X) = \frac{182-147}{12} = \frac{35}{12}$

b) i) $P(\text{red}) = \frac{1}{2}$, $P(\text{blue}) = \frac{1}{3}$, $P(\text{green}) = \frac{1}{6}$

The desired probability is the multinomial

prob: $\frac{3!}{1!\,1!\,1!}\left(\frac{1}{2}\right)^1\left(\frac{1}{3}\right)^1\left(\frac{1}{6}\right)^1 = \frac{1}{6}$

ii) From (a), if S denotes the sum then

$E(S) = 3 \times 3.5 = 10.5$ and

$Var(S) = 3 \times \frac{35}{12} = \frac{35}{4} = 8.75.$

$P(S \leq 6) = P(S \leq 6.5) \approx P(Y \leq 6.5)$

where $Y \sim N(10.5, 8.75)$

$$\cdots = P\left(\frac{Y - 10.5}{\sqrt{8.75}} \leq \frac{6.5 - 10.5}{\sqrt{8.75}} \right)$$

$$\approx P(Z \leq -1.35)$$

$$= 1 - \Phi(1.35)$$

$$\approx 1 - 0.9115 \quad \text{from tables}$$

$$= 0.0885$$

Note: we could perhaps improve the approximation by writing

$$P(S \leq 6) = P(3 \leq S \leq 6)$$

$$= P(2.5 \leq S \leq 6.5)$$

$$\approx P(2.5 \leq Y \leq 6.5)$$

$$= \ldots \text{ etc.}$$

(iii) Write $p(x,y,z) = P(\text{first} = x, \text{second} = y, \text{third} = z)$

$$= \frac{1}{6^3} = \frac{1}{216}$$

for all possible $x, y, z = 1, 2, \ldots, 6$.

$$P(S \le 6) = P(S=3) + P(S=4) + P(S=5) + P(S=6).$$

$$P(S=3) = p(1,1,1) = \frac{1}{216}.$$

$$P(S=4) = p(1,1,2) + p(1,2,1) + p(2,1,1) = \frac{3}{216}$$

$$P(S=5) = p(1,1,3) + p(1,3,1) + p(3,1,1)$$
$$+ p(1,2,2) + p(2,1,2) + p(2,2,1) = \frac{6}{216}$$

$$P(S=6) = p(1,1,4) + p(1,4,1) + p(4,1,1)$$
$$+ p(1,2,3) + \cdots\cdots + p(3,2,1)$$
$$+ p(2,2,2)$$
$$= \frac{10}{216}$$

So $P(S \le 6) = \frac{20}{216} \approx 0.093.$

4. a)

	obs.	exp.
A	22	27
B	14	9

(Pearson's) χ^2-statistic $= \dfrac{(22-27)^2}{27} + \dfrac{(14-9)^2}{9}$

$$= 25\left[\frac{1}{27} + \frac{1}{9}\right] = 25\left(\frac{4}{27}\right) = \frac{100}{27}$$

$$\simeq 3.704$$

p-value $= P(\chi^2_1 \geq 3.704)$

From χ^2 tables, since 3.704 is between 2.706 and 3.841 the p-value is between 0.05 and 0.10.

[Note: $P(\chi^2_1 \geq 3.704) = 2P(Z \geq \sqrt{3.704} \simeq 1.92)$

$$\simeq 2[1 - \Phi(1.92)]$$
$$\simeq 0.0548.$$]

b) The χ^2 statistic is.

$$\frac{(x-np_0)^2}{np_0} + \frac{\left[(\cancel{n}-x) - n(\cancel{1}-p_0)\right]^2}{n(1-p_0)}$$

$$= (x-np_0)^2\left[\frac{1}{np_0} + \frac{1}{n(1-p_0)}\right] = \frac{(x-np_0)^2}{np_0(1-p_0)}$$

So the χ^2-test p-value $\geq \alpha$ if and only if

$$\frac{(x-np_0)^2}{np_0(1-p_0)} \leq c \quad \text{where} \quad P(\chi^2_1 > c) = \alpha.$$

But this value of c also satisfies.

$$2P(Z > \sqrt{c}) = \alpha$$

$$P(Z > \sqrt{c}) = \frac{\alpha}{2}.$$

$$\sqrt{c} = z_{\alpha/2}.$$

So the χ^2-test p-value $\geq \alpha$

$$\iff \frac{(x-np_0)^2}{np_0(1-p_0)} \leq z_{\alpha/2}^2$$

$$\iff \frac{|x-np_0|}{\sqrt{np_0(1-p_0)}} \leq z_{\alpha/2}$$

$$\iff \frac{\sqrt{n}\,|\hat{p}-p_0|}{\sqrt{p_0(1-p_0)}} \leq z_{\alpha/2}$$

$\iff$ p_0 is in $100(1-\alpha)\%$ Wilson C.I.