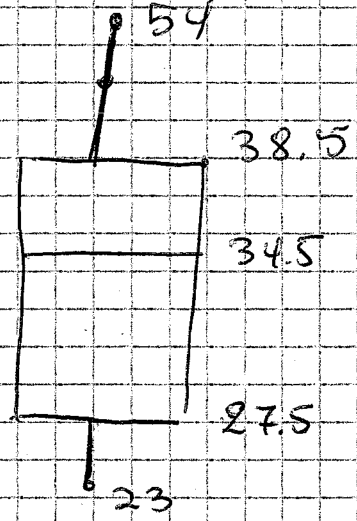
1a) Let X_1, X_2, X_3 be the three salaries.

Then $\dfrac{X_1+X_2+X_3}{3} = \bar{X} = 125\,000 \Rightarrow$

2.0

$X_1+X_2+X_3 = 375\,000$ which is higher than 400,000 \$, so No it is not possible.

1b) 23, 24, 26, 29, 33, 34 35, 37, 38, 39, 48, 54



Mean = 34.17

Median = $\dfrac{34+35}{2} = \underline{\underline{34.5}}$

25% quartile: $\dfrac{26+29}{2} = \dfrac{55}{2} = \underline{\underline{27.5}}$

75% quartile: $\dfrac{38+39}{2} = \underline{\underline{38.5}}$

min = $\underline{\underline{23}}$

max = $\underline{\underline{54}}$

1

$IQR = 38.5 - 27.5 = \underline{11}$

A possible outlier is $> Q_3 + 1.5\,IQR = 38.5 + 1.5 * 11 =$

$= 55$, so 54 is not a possible outlier.

1

No outliers

2) $P(A)=0.25 \quad P(B)=0.4$

a) $P(A\cap B) \overset{\text{indep}}{=} P(A)\cdot P(B) = 0.25 * 0.4 = 0.1$ (1)

b) $P(A|B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{0.1}{0.4} = 0.25$ (1)

OR

$= \dfrac{P(A\cap B)}{P(B)} = \dfrac{P(A)\cdot P(B)}{P(B)} = P(A) = 0.25$

c) $P(A\cup B) = P(A)+P(B)-P(A\cap B) = 0.25+0.4-0.1 = \underline{\underline{0.55}}$

(2)

3) $f(x) = \dfrac{k}{2^x}, \quad x=0,1,2,3,4$

a) $\sum_{x=0}^{4} f(x) = 1$ and $f(x)\geq 0 \quad \forall x$

$1 = \dfrac{k}{2^0}+\dfrac{k}{2}+\dfrac{k}{2^2}+\dfrac{k}{2^3}+\dfrac{k}{2^4} = \dfrac{k}{1}+\dfrac{k}{2}+\dfrac{k}{4}+\dfrac{k}{8}+\dfrac{k}{16}$

$= \dfrac{16k+8k+4k+2k+k}{16} = \dfrac{31k}{16} = 1 \Rightarrow \boxed{k=\dfrac{16}{31}}$ (1)

b) $EX = \sum_{x=0}^{4} x\cdot P(X=x) = 0f(0)+1\cdot f(1)+2\cdot f(2)+3\cdot f(3)$

$+4f(4) = \dfrac{16}{31}\cdot\left\{\dfrac{1}{2}+2\cdot\dfrac{1}{2^2}+3\cdot\dfrac{1}{2^3}+4\cdot\dfrac{1}{2^4}\right\} =$

$= \dfrac{16}{31}\cdot\left\{\dfrac{1}{2}+\dfrac{1}{2}+\dfrac{3}{8}+\dfrac{4}{4}\right\} = \dfrac{16}{31}\cdot\left\{\dfrac{4+4+3+2}{8}\right\} =$

$= \dfrac{\cancel{16}^{2}}{31}\cdot\dfrac{13}{\cancel{8}} = 26/31.$

(1.5)

c) $P(\bar{X} \le 2) = P(\bar{X}=0) + P(\bar{X}=1) + P(\bar{X}=2) =$

$$= \frac{16}{31} \cdot \left(\frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} \right) = \frac{16}{31} \cdot \frac{7}{4} = \frac{28}{31}$$ 1

d) $P(\bar{X} \in B) = P(X=1, X=3) = \frac{16}{31}\left(\frac{1}{2} + \frac{1}{2^3}\right) = \frac{16}{31}\left(\frac{4}{8} + \frac{1}{8}\right) = \frac{16}{31} \cdot \frac{5}{8} = \frac{10}{31}$

$P(\bar{X}=1 \mid X \in B) = \frac{P(X=1 \cap X \in B)}{P(X \in B)} = \frac{P(X=1)}{P(X \in B)} = \frac{\frac{16}{31} \cdot \frac{1}{2}}{10/31} = \frac{8/31}{10/31} = 4/5$ 2P

4 a) A Bernoulli trial is any experiment that has only 2 possible outcomes, failure or success 0.5

b) $Var\, Y = Var(2X+3) = 4 \cdot Var\, X \neq Var\, X$ 0.5

c) $\bar{Y}$ or (median) 0.5

d) Point estimate is the value of the point estimator after sample has been taken 0.5

e) two-sided 0.5

f) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.45 + 0.7 - 0.1 = 1.05 > 1$

impossible 0.5

5 a) A good model for Y is the binomial. Each test can be thought of as a Bernoulli trial with 2 possible outcomes False or Not-false and constant probability of getting a False $= 0.0015$ from one test to another.

(2)

$Y \sim B(10000, 0.0015)$.

b) Since $n = 10.000$ is really big we can approximate the Probability $P(Y \leq 20)$ by approximating the Binomial distr by a normal distr.

$Y \sim N(n.p, n.p(1-p)) = N(15, 15 * (1-0.0015)) =$

$\sim N(15, 15)$. (1)

Then $P(Y \leq 20) = P\left(\frac{Y-15+0.5}{\sqrt{15}} \leq \frac{20-15+0.5}{\sqrt{15}}\right) =$

$= P\left(Z \leq \frac{5.5}{\sqrt{15}}\right) \approx P(Z \leq 1.42) = \underline{\underline{0.9222}}$ (1).

$\underline{\text{OR}}$ It could be approximated using a Poisson distr since $n = 10000 \geq 100$, $np = 15 \leq 20$ and $p < 0.01$

$\lambda = n\,p = 15$

$P(Y \leq 15) = \sum_{x=0}^{20} \frac{15^x}{x!} e^{-15} = 0.917$

6) ~~a)~~ These are two independent populations and estimated variances different $S_1^2 = 3.16$, $S_2^2 = 12.13$

a) $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 < 0$

$$t = \frac{\bar{Y}_1 - \bar{Y}_2 - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{64.68 - 66.28}{\sqrt{\frac{3.16}{10} + \frac{12.13}{10}}} = \frac{-1.6}{\cancel{2.09}\ 1.2365} = \cancel{-0.7654}\ -1.294$$

1

degrees of freedom

$$\nu = \left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2 \Big/ \left[\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}\right] =$$

$$= (1.529)^2 \Big/ \left(\frac{0.316^2}{9} + \frac{1.213^2}{9}\right) = 2.3378 \big/ (0.0111 + 0.1635)$$

$$= \frac{2.3378}{0.1746} = 13.39$$

$\nu = 13$

$t_{0.025,13} = t_{0.975,13} = 2.1604$

Reject H_0 if $t_0 < -t_{0.975,13}$

But $-1.294 > -2.1604$

1

So we don't have enough evidence to reject H_0

(~~7~~ 6 b) The two populations are independent

$Y_{1,j} = \mu_1 + \varepsilon_{1,j}$, $\varepsilon_{1,j} \sim N(0, \sigma_1^2)$

$Y_{2,j} = \mu_2 + \varepsilon_{2,j}$ $\varepsilon_{2,j} \sim N(0, \sigma_2^2)$ 1

σ_1^2, σ_2^2 unknown and $\sigma_1^2 \neq \sigma_2^2$.

c) A 95% C.I is of the form

$$\mu_1 - \mu_2 \in \left(\bar{y}_1 - \bar{y}_2 \pm t_{0.025,13} * S\right) =$$

$$= \left(-1.6 \pm 2.1604 * 1.2365\right) =$$ 1

$$= (-4.2713, 1.0713)$$

$$S = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{3.16}{10} + \frac{12.13}{10}} = \underline{\underline{1.2365}}$$

d) The C.I contains 0 so it does not contradict a). 0.5

7)

	Low	Aver	High	Total
Low	6 13.68	12 23.16	32 13.16	50
Aver	33 30.65	61 51.9	18 29.5	112
High	13 7.66	15 12.97	0 7.37	28
Total	52	88	50	190

H_0: Fidelity and Selectivity are Independent

H_a: they are not

$$\chi_0^2 = \sum_{ij} \frac{(O_{ij}-E_{ij})^2}{E_{ij}} = \frac{(6-13.68)^2}{13.68} + \frac{(12-23.16)^2}{23.16} + \frac{(32-13.16)^2}{13.16}$$

$$+ \frac{(33-30.65)^2}{30.65} + \frac{(61-51.9)^2}{51.9} + \frac{(18-29.5)^2}{29.5} + \frac{(13-7.66)^2}{7.66} +$$

$$+ \frac{(15-12.97)^2}{12.97} + \frac{(0-7.37)^2}{7.37} = 54.3299 = \underline{\underline{54.33}} \quad 1$$

Reject H_0 if $\chi_0^2 > \chi^2_{\alpha,(r-1)\cdot(c-1)}$

$\chi^2_{0.01,4} = 0.297$ 0.5

$\boxed{\text{So reject } H_0}$ 1

8) a) $b = \frac{S_{xy}}{S_{xx}} = \frac{230}{854.92} = 0.269$

$a = \bar{y} - b\,\bar{x} = 11 - 0.269 * 44.42 = -0.9490$

$y = \alpha + \beta x + \varepsilon$ or $\hat{y} = a + bx \Rightarrow$

$\boxed{\hat{y} = -0.949 + 0.269 \cdot x}$ 1

b) $\hat{y}$ when $x = 40$ is $\hat{y} = -0.949 + 0.269 * 40 =$

$= 9.811$ 1

c) A 99% CI for mean moisture content at $x=40$ is given by:

$$Y \in \left[\hat{y}\big|_{x=x_0} \pm t_{0.005,10} \cdot \hat{\sigma} \cdot \sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}} \right]$$

where $\hat{y}\big|_{x=x_0} = 9.811$, $t_{0.005} = 3.1693$

where $$\hat{\sigma}^2 = \frac{S_{yy} - (S_{xy})^2/S_{xx}}{n-2} =$$

$$= \frac{74 - 230^2/854.92}{10} = 1.2123 \Rightarrow$$

$$\boxed{\hat{\sigma} = 1.101} \quad 1$$

Hence $$Y \in \left[9.811 \pm 3.1693 * 1.101 \cdot \sqrt{\frac{1}{12} + \frac{(40-44.42)^2}{854.92}}\right]$$

$$= [9.811 \pm 3.1693 * 1.101 * 0.3259] =$$

$$= [8.6738, 10.9482]. \quad 1$$