Question 1.

Marks & seen/unseen

Parts

seen ⇓

seen sim. ⇓

(i) (c).

(ii) (a)

Let R denote relevant, W denote chosen words present

$P(R|W) = \dfrac{P(W|R)P(R)}{P(W)} = \dfrac{1 * 0.01}{1 * 0.01 + 0.1 * 0.99} = 0.092$

(iii) (a). $np = 3$ and $np(1-p) = 2$, so $1 - p = 2/3$ ie $p = 1/3$ and $n = 9$.

(iv) (c).

(v) (f).

Denoting T the lifetime of the system, and T_1 and T_2 as the two independent lifetimes with exponential distributions with parameter λ,

$P(T < t) = P(T_1 < t \cap T_2 < t) = P(T_1 < t)P(T_2 < t) = (1 - e^{-\lambda t})^2$

So the survivor function, $P(T > t) = 1 - (1 - e^{-\lambda t})^2 = 2e^{-\lambda t} - e^{-2\lambda t}$

Each 4 marks

Question

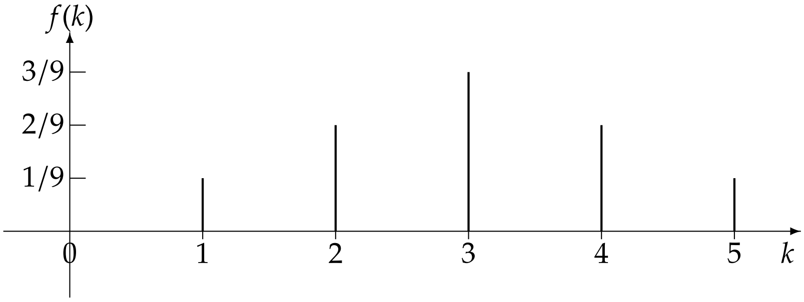
2.

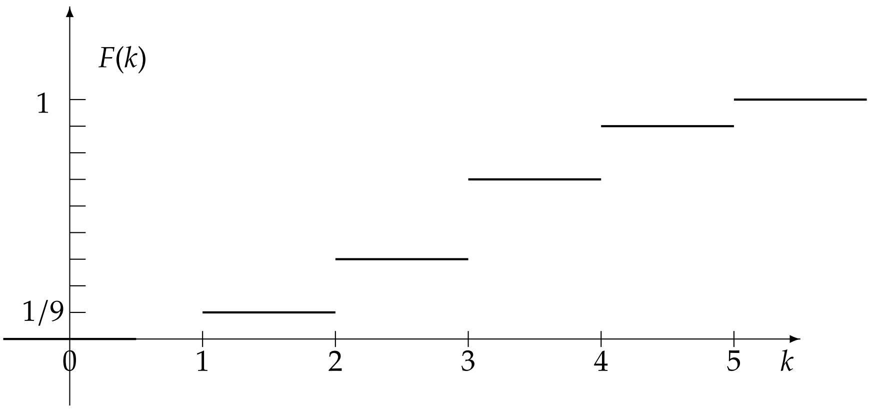
Marks & seen/unseen

Parts

seen sim. ⇓

(i) (a)





5 marks

(b) $E(X) = 1 * 1/9 + 2 * 2/9 + 3 * 3/9 + 4 * 2/9 + 5 * 1/9 = 27/9 = 3$ — 2 marks

(ii) (a) Table of expected counts

k	frequency O_k	expected freq E_k
1	25	100*1/9=11.1
2	35	100*2/9=22.2
3	21	100*3/9=33.3
4	13	100*2/9=22.2
5	6	100*1/9=11.1

3 marks

(b) H_0 : Observed counts equal expected counts — 2 marks

(c)

$$X^2 = \sum_{i=1}^{5} \frac{(O_k - E_k)^2}{E_k} \sim \chi^2_4$$

under the null hypothesis.

$$X^2 = \frac{(25-11.1)^2}{11.1} + \frac{(35-22.2)^2}{22.2} + \frac{(21-33.3)^2}{33.3} + \frac{(13-22.2)^2}{22.2} + \frac{(6-11.1)^2}{11.1} = 35.5$$

The p-value is the probability of observing these data, or data more extreme than these, if the null hypothesis is true. This is very small, <0.005 from χ^2 table. — 6 marks

Question	2.	Marks & seen/unseen
Parts	(ii)(d) Reject the null hypothesis at both 5% and 1% significance. The model does not fit the data well.	seen sim. ⇓ 2 marks

Question 3.

Marks & seen/unseen

Parts

seen ⇓

(i) (a) likelihood function for (μ, σ^2) for all of the data is

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f_X(x_i)$$
$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}\right\}.$$

The log likelihood is

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}.$$

For MLE for μ, take partial derivative wrt μ and set this equal to zero and solve for $\hat{\mu}$.

$$0 = \frac{\partial}{\partial \mu}\ell(\mu, \sigma^2)|_{\mu=\hat{\mu}} = \frac{\sum_{i=1}^{n}(x_i - \hat{\mu})}{\sigma^2}$$
$$\iff 0 = \sum_{i=1}^{n}(x_i - \hat{\mu}) = \sum_{i=1}^{n} x_i - n\hat{\mu} \iff \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}.$$

To check this is a maximum, look at the second derivative.

$$\frac{\partial^2}{\partial \mu^2}\ell(\mu, \sigma^2) = \frac{-n}{\sigma^2},$$

which is negative everywhere, so $\bar{x}$ is the MLE for μ, independently from the value of σ^2. 7 marks

(b)

$$\mathrm{E}(\overline{X}) = \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{1}{n}n\mu = \mu.$$

Hence $\overline{X}$ is an unbiased estimator of μ. 3 marks

(ii) (a) Sample mean $\bar{x} = 1.175$, $s = 1.514$ where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$. seen sim. ⇓

(b) Assuming the time differences are normally distributed. Use a t distribution with 7 df for the confidence interval. 4 marks

$$1.175 \pm t_{7,0.975} * 1.514/\sqrt{8} = 1.175 \pm 2.36 * 1.514/\sqrt{8} = (-0.088, 2.438)$$

3 marks

(c) The 95% confidence interval includes 0, so the p-value for a two-sided test is >0.05. At 5% significance, we have insufficient evidence to reject the null hypothesis that the population mean time difference is zero. 3 marks

Question 4.

Marks & seen/unseen

Parts

seen ⇓

(i) For f to be a valid probability density function,

I. $f(x) \geq 0, \forall x \in \mathbb{R}$;

II. $\int_{x=-\infty}^{\infty} f(x)dx = 1.$

2 marks

seen sim. ⇓

(ii) From part (i), $\int_{x=0}^{1} f(x)\mathrm{d}x = 1$. So $\int_{x=0}^{1} a + bx^2\mathrm{d}x = [ax + \frac{bx^3}{3}]_0^1 = a + b/3 = 1$

And $\mathrm{E}(X) = \frac{3}{8} \Rightarrow \int_{x=0}^{1} x(a + bx^2)\mathrm{d}x = \frac{3}{8} \Rightarrow [\frac{ax^2}{2} + \frac{bx^4}{4}]_0^1 = \frac{a}{2} + \frac{b}{4} = \frac{3}{8}$

Solving gives $a = 3/2$ and $b = -3/2$

4 marks

(iii) $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$

$\mathrm{E}(X^2) = \frac{3}{2}\int_{x=0}^{1} x^2(1 - x^2)\mathrm{d}x = \frac{3}{2}[\frac{x^3}{3} - \frac{x^5}{5}]_0^1 = \frac{3}{2}(\frac{1}{3} - \frac{1}{5}) = \frac{3}{15} = \frac{1}{5}$

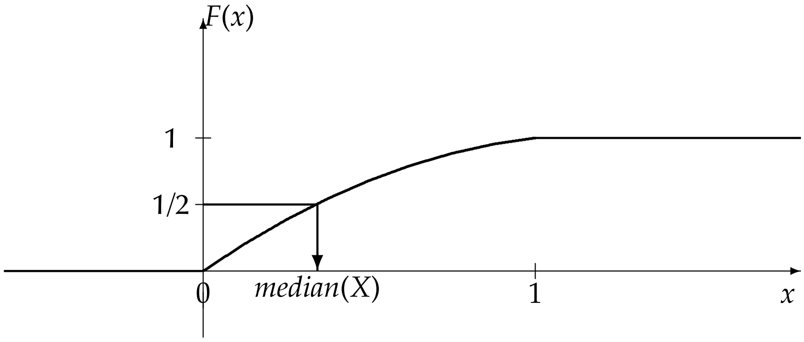
So $\mathrm{Var}(X) = \frac{1}{5} - (\frac{3}{8})^2 = \frac{64 - 45}{5 * 64} = \frac{19}{320} = 0.059$

4 marks

(iv) For $0 \leq x \leq 1$

$$F(x) = \int_{u=0}^{x} f(u)\mathrm{d}u = \int_{u=0}^{x} \frac{3}{2}(1 - u^2)\mathrm{d}u = \frac{3}{2}[u - \frac{u^3}{3}]_0^x = \frac{3x - x^3}{2}$$

For $x < 0$, $F(x) = 0$ and for $x > 1$, $F(x) = 1$



4 marks

(v) (a) $P(Y \leq y) = P(3X - 1 \leq y) = P(3X \leq y + 1) = P(X \leq \frac{y+1}{3}) = F_X(\frac{y+1}{3}) = \frac{1}{2} * (3 * (\frac{y+1}{3}) - (\frac{y+1}{3})^3) = \frac{y+1}{2} - \frac{1}{54} * (y + 1)^3$

Then, $f_Y(y) = \frac{dF_Y}{dy} = \frac{1}{2} - \frac{1}{18}(y + 1)^2$ where $-1 \leq y \leq 2$.

4 marks

(b) $\mathrm{E}(Y) = \mathrm{E}(3X - 1) = 3\mathrm{E}(X) - 1 = 3 * 3/8 - 1 = 1/8$.

$\mathrm{Var}(Y) = \mathrm{Var}(3X - 1) = 3^2 * \mathrm{Var}(X) = 9 * 19/320 = \frac{171}{320} = 0.53$.

2 marks