Q1.

(i) Two events $A$ & $B$ s.t. $AB = \phi$.

$$P[\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P[E_i]$$

Proof: Given $\{A_i\}$ for $i=1, \cdots \infty$ s.t.
$A_i = E_i$ for $i = 1, \cdots, n$
$A_i = \phi$ for $i > n$.

Since the collection $\{A_i\}$ consists of M.E. events, so axiom(3) applies:

$$\bigcup_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{n} E_i) \cup A_{i+1} \cup A_{i+2} \cup \cdots$$

Note $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} E_i \Rightarrow P[\bigcup_{i=1}^{\infty} A_i] = P[\bigcup_{i=1}^{n} E_i]$

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(A_i)$$

$$= \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{n} P(E_i)$$

$$\Rightarrow P[\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P[E_i]$$

(ii) We say event $A$ is indep of $B$ if

$$P(A|B) = P(A)$$

<u>Proof:</u> $P(A^c \cap B) = P[B \setminus (AB)] = P(B) - P(AB)$
$= P(B) - P(A)P(B)$
$= P(B) \cdot P(A^c)$

$\Rightarrow A^c$ indep of $B$.

$$P(A^c \cap B^c) = P[(A \cup B)^c] = 1 - P(A \cup B)$$
$$= 1 - P(A) - P(B) + P(AB)$$
$$= 1 - P(A) - P(B) + P(A)P(B).$$
$$= (1 - P(A)) - P(B)(1 - P(A))$$
$$= (1 - P(A))(1 - P(B)) = P(A^c)P(B^c)$$

$\Rightarrow A^c$ & $B^c$ are indep.

(iii) Each urn contains the same number $n-1$ of balls, there are in total $n(n-1)$ balls. The second ball picked is equally likely. Denote as $C_i$ the colour of the $i$th ball picked. $\Rightarrow P[C_2 = M] = \frac{1}{2}$.

(iv). $P[C_2=M \mid C_1=M] = \dfrac{P[C_1, C_2=M]}{P[C_1=M]} = 2P[C_1, C_2=M]$

in particular, $P[C_1, C_2=M] = \dfrac{\sum_{r=1}^{n}(n-r)(n-r-1)}{n(n-1)(n-2)}$

$\sum_{r=1}^{n}(n-r)^2-(n-r)$, variable chang, $n-r=k$.

$= \sum_{k=0}^{n-1}(k^2-k) = \sum_{k=1}^{n-1}k^2 - \sum_{k=1}^{n-1}k$.

$2(n-1)+1$

$2n-1$.

$= \dfrac{(n-1)(n)(2n-1)}{6} - \dfrac{3n(n-1)}{6}$

$P[C_1, C_2=M] = \dfrac{\frac{\cancel{n(n-1)}(2n-1)-3\cancel{n(n-1)}}{6}}{\cancel{n(n-1)}(n-2)}$

$= \dfrac{2n-1-3}{6(n-2)} = \dfrac{2(n-2)}{6(n-2)} = \dfrac{1}{3}$

$\Rightarrow P[C_2=M \mid C_1=M] = \dfrac{2}{3}$.

Q 2.

Assume that each of the six orderings of the car and goats are equally likely. Let $C_i$ be the event that the $i^{th}$ door conceals the car, $G$ the event that you see a goat, $B$ the event that you see $B$

(i) $$P[C_3 | G] = \frac{P[C_3 \cap G | C_1] P[C_1] + P[C_3 \cap G | C_1^c) P(C_1^c)}{P[G | C_1] P(C_1) + P(G | C_1^c) P(C_1^c)}$$

$$= \frac{0 \times \frac{1}{3} + 1 \times \frac{2}{3}}{1 \times \frac{1}{3} + 1 \times \frac{2}{3}} = \frac{2}{3}.$$

(ii)

$$P(C_3 | B) = \frac{P(C_3 \cap B)}{P(B)}$$

$$= \frac{P(C_3 \cap B | C_1) P(C_1) + P(C_3 \cap B | C_1^c) P(C_1^c)}{P(B | C_1) P(C_1) + P(B | C_1^c) P(C_1^c)}.$$

Note $P(C_3 \cap B | C_1^c) = P(B | C_1^c) = \frac{1}{2}.$

$$= \frac{0 + \frac{1}{2} \times \frac{2}{3}}{b \times \frac{1}{3} + \frac{1}{2} \times \frac{2}{3}} = \frac{\frac{1}{3}}{\frac{b}{3} + \frac{1}{3}} = \frac{1}{1+b}$$

(iii). This time

$$P(C_3 \cap G \mid C_1^c) = P(G \mid C_1^c) = \frac{1}{2},$$

Presenter needs to randomly select G out of two gates, with equal likelihood. Hence $\frac{1}{2}$.

$$\Rightarrow p = \frac{0 + \frac{1}{2} \times \frac{2}{3}}{1 \cdot \frac{1}{3} + \frac{1}{2} \times \frac{2}{3}} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}.$$

(iv) Let $\alpha \in [\frac{1}{2}, \frac{2}{3}]$, and suppose the presenter possesses a coin which falls with heads upwards with prob. $\beta = 6\alpha - 3$. He flips the coin before the show, and adopts strategy (i) iff the coin shows heads, and otherwise strategy (iii). The prob. in question is now $\frac{2}{3}\beta + \frac{1}{2}(1-\beta) = \alpha$. You never lose by swapping, but whether you gain depends on the presenter's protocol.

Q 3.

(i) $P[Y < y \mid Y < 4] = \dfrac{P[Y < y \cap Y < 4]}{P[Y < 4]}$

$P[Y < 4] = \int_0^4 f(y)\, dy = 0.352.$

$P[Y < y \cap Y < 4] = P[Y < y]$ for $0 \le y \le 4$

$= \int_0^y \dfrac{60y - 6y^2}{1000}\, dy$

$= \dfrac{1}{1000}\left[30y^2 - 2y^3\right]\Big|_0^y = \dfrac{30y^2 - 2y^3}{1000}$

$\Rightarrow \boxed{P[Y < y \mid Y < 4] = \left(\dfrac{30y^2 - 2y^3}{1000}\right) \times \dfrac{125}{44}}$

For $y > 4$, $\Rightarrow \boxed{P[Y < y \mid Y < 4] = 1.}$

(ii) $f_{Y \mid Y<4}(y \mid y < 4) = \dfrac{d\,P[Y<y \mid Y<4]}{dy} = \dfrac{15}{88}y - \dfrac{3}{176}y^2$

(iii) $E[Y \mid Y<4] = \int_0^4 y \cdot f(y \mid y<4)\, dy =$

$$= \int_0^4 y \cdot \frac{15}{88} y - \frac{3}{176} y^3 \, dy$$

$$= \frac{1}{88} \int_0^4 15y^2 - \frac{3}{2} y^3 \, dy = \frac{1}{88} \left[ 5y^3 - \frac{3}{8} y^4 \right] \Big|_0^4$$

$$= \frac{1}{\cancel{88}_{11}} \left[ 5 \times \cancel{64}^{8} - \frac{3}{\cancel{8}} \times \cancel{256}^{32} \right] = \frac{28}{11}$$

$$E[Y^2 \mid Y<4] = \int_0^4 y^2 \cdot f(y \mid y<4)\, dy = \frac{408}{55}$$

$$Var[Y \mid Y<4] = \frac{408}{55} - \left(\frac{28}{11}\right)^2 = 0.93886$$

(iv) $P[Y<y \mid 1<Y<4] = \dfrac{P[Y<y \cap (1<Y<4)]}{P[1<Y<4]}$

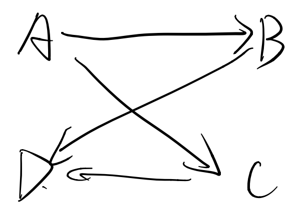
$$P[1<Y<4] = \frac{44}{125} - P[Y<1] = \frac{44}{125} - \frac{7}{250} = \frac{81}{250}$$

For $y<1 \Rightarrow 0$

$1 \leq y \leq 4 \Rightarrow \left(\dfrac{30y^2 - 2y^3}{1000}\right)\left(\dfrac{250}{81}\right)$

$y>4 \Rightarrow 1$

Q4.




Use $A \leftrightarrow D$ for information is transferred from A to D., EF for the event that there is not fog between E and F.

(i) $P[A \leftrightarrow D \mid AD^c] = P[A \leftrightarrow D \mid AD^c \cap BC^c]\, p$
$\qquad + P[A \leftrightarrow D \mid AD^c \cap BC](1-p)$

$= \{1-(1-(1-p)^2)^2\}\, p + (1-p^2)^2(1-p)$

(ii) $P[A \leftrightarrow D \mid BC^c] = P[A \leftrightarrow D \mid AD^c \cap BC^c)\, p$
$\qquad + P[A \leftrightarrow D \mid BC^c \cap AD)(1-p)$

$= \{1-(1-(1-p)^2)^2\}\, p + (1-p)$

(iii) $P[A \leftrightarrow D \mid AB^c] = P[A \leftrightarrow D \mid AB^c \cap AD^c)p + P[A \leftrightarrow D \mid AB^c \cap AD]p$

$= (1-p)\{1-p(1-(1-p)^2)\}p + \{1-p)$

(iv) $P[A \leftrightarrow D] = P[A \leftrightarrow D \mid AD^c)p + P[A \leftrightarrow D \mid AD](1-p)$

$= \{1-(1-(1-p)^2)^2\}p^2 + (1-p^2)^2 p(1-p) + (1-p)$

Q5.

(i). For a Random variable $X$ & some $f^n$. $g(x) \geq 0$ and constant $c$.

$$P[g(x) \geq c] \leq \frac{E[g(x)]}{c}$$

<u>Proof:</u> $E[g(x)] = \int_{\text{all } x} g(x) f(x)\, dx$

$$= \underbrace{\int_{\substack{\text{all } x \\ \text{s.t. } g(x) \geq c}} g(x) f(x)\, dx}_{①} + \underbrace{\int_{\substack{\text{all } x \\ \text{s.t. } g(x) < c}} g(x) f(x)\, dx}_{②}$$

② $\geq 0$ since $g(x) \geq 0$.

$$\Rightarrow E[g(x)] \geq \int_{\substack{\text{all } x \\ \text{s.t. } g(x) \geq c}} g(x) f(x)\, dx \geq \int_{\substack{\text{all } x \\ \text{s.t. } g(x) \geq c}} c\, f(x)\, dx$$

$$= c \int_{\substack{\text{all } x \\ \text{s.t. } g(x) \geq c}} f(x)\, dx = c\, P[g(x) \geq c]$$

$$\Rightarrow P[g(x) \geq c] \leq \frac{E[g(x)]}{c}$$

ii) Suppose we have some R.V. $X$ for which $Var(X) = \sigma^2$ exists.
Then $P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$,
where $\mu = E(X)$, $k \in \mathbb{Z}^+$ usually $k \geq 1$.

<u>Proof:</u> Choose $g(x) = (x-\mu)^2$, $c = k^2\sigma^2$

$$\Rightarrow P[(X-\mu)^2 \geq k^2\sigma^2] \leq \frac{E[g(x)]}{c} = \frac{\sigma^2}{k^2\sigma^2}$$

$$\Rightarrow P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$$

iii) Consider a sequence of R.V. $\{X_n\}$. We say that the sequence converges in prob to the value $c$ if for every $\varepsilon > 0$

$$\lim_{n\to\infty} Prob[|X_n - c| < \varepsilon] = 1$$

(iv) $X_1, ---,$ are iid distributed with a finite first moment $\mu$, then for the sequence of averages $\bar{X}_n = \frac{1}{n} \sum X_i$, we have

$$\bar{X}_n \xrightarrow{P} \mu.$$

<u>Proof</u>: Assume $X_i$ has a finite variance $\sigma^2$.

For $\bar{X}_n$, $E[\bar{X}_n] = \mu$, $Var(\bar{X}_n) = \frac{\sigma^2}{n}$

For any $\varepsilon > 0$, choose $k$ s.t. $\frac{k\sigma}{\sqrt{n}} = \varepsilon$, i.e. $k = \frac{\varepsilon\sqrt{n}}{\sigma}$

$\Rightarrow P[\,|\bar{X}_n - \mu| < \varepsilon\,] \geq 1 - \frac{\sigma^2}{n\varepsilon^2}$

$\Rightarrow \lim_{n\to\infty} P[\,|\bar{X}_n - \mu| < \varepsilon\,] = 1.$ #

Question 6

(i) $P[Y=k] = \int_0^1 P[Y=k \mid X=x]\, f_X(x)\, dx$

$$= \int_0^1 \binom{n}{k} x^k (1-x)^{n-k} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\, dx$$

$$= \binom{n}{k} \frac{B(a+k, n-k+b)}{B(a,b)}$$

(ii) $E(Y) = E_X[E_Y[Y \mid X]] = \int_0^1 E_Y(Y \mid X=x) f_X(x)\, dx$

$$= \frac{na}{a+b}$$

(iii) $E(Y^2) = E_X[E_Y[Y^2 \mid X]] = \int_0^1 E_Y(Y^2 \mid X=x) f_X(x)\, dx$

$$= \int_0^1 [nx(1-x) + n^2x^2] \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\, dx$$

$$= \int_0^1 \frac{n x^a (1-x)^b}{B(a,b)} + \frac{n^2 x^{a+1}(1-x)^{b-1}}{B(a,b)}\, dx$$

$$= \frac{n}{B(a,b)} \cdot B(a+1, b+1) + n^2 \cdot \frac{B(a+2, b)}{B(a,b)}$$

$$= \frac{abn}{(a+b+1)(a+b)} + \frac{(a+1)\, a\, n^2}{(a+b+1)(a+b)}$$

$$= \frac{abn + (a+1)an^2}{(a+b)(a+b+1)} \qquad \frac{a^2n^2}{(a+b)^2}$$

$$Var(Y) = E(Y^2) - E(Y)^2$$

$$= \frac{abn(a+b) + (a+b)(a+1)an^2 - a^2n^2(a+b+1)}{(a+b)^2(a+b+1)}$$

$$-a^2n^2$$

$$= \frac{\cancel{a^2bn} + ab^2n + \cancel{a^3n^2} + \cancel{a^2n^2} + a^2bn^2 + abn^2 - \cancel{a^3n^2} - \cancel{a^2bn}}{(a+b)^2(a+b+1)}$$

$$\boxed{= \frac{abn(a+b+n)}{(a+b+1)(a+b)^2}}$$

(iv) $X$ uniform $\Rightarrow$ $a = b = 1$.

Hence $P[Y=k] = \binom{n}{k} \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} = \frac{1}{n+1}$,

$0 \le k \le n$.

$\Rightarrow$ $Y$ is uniformly distributed.

Question 7.

(i)

Eigenvalues :

$$|S - l| = l^2 - (a+b)l + (ab - c^2)$$

Eigenvalues satisfy $l = \dfrac{(a+b) \pm \sqrt{(a+b)^2 - 4(ab-c^2)}}{2}$

2 marks

For eigenvector $x = (x_1, x_2)'$, $Sx = lx$ and $x_1^2 + x_2^2 = 1$.

$$x_1 = -cx_2/(a-l), \quad x_2 = 1/\sqrt{1 + c^2/(a-l)^2}$$

If $\Gamma$ is the matrix with the 2 eigenvectors as columns, the principal components of X are given by

$$Z = \Gamma'(X - \mu).$$

Variance of first and second principal components are given by the eigenvalues.

(ii)

$$\phi_Y(t) = E[e^{is^T Y}] = E[e^{is^T AX}] = E[e^{i(A^T s)^T X}]$$

$$= e^{i(A^T s)^T \mu - \frac{1}{2}(A^T s)^T \Sigma (A^T s)}$$

$$= e^{is^T(A\mu) - \frac{1}{2} s^T (A \Sigma A^T) s}$$

$\Rightarrow Y \sim N_q(A\mu, A\Sigma A^T)$.

(iii) $Z = W^{-1}(X-\mu)$, by (ii) $Z$ is normal.

$E(Z) = 0$. $Var(Z) = W^{-1} Var(X) (W^{-1})^T$

$$= W^{-1} \cdot \Sigma (W^{-1})^T$$

$$= W^{-1} \cdot W \cdot W \cdot (W^{-1})^T$$

$$= I \cdot (W^{-1} W^T)^T = I.$$

$\Rightarrow Z \sim N_p(0, I)$.

Question 8.

(i).

$$G_{Y_1}(t) = \sum_{r=0}^{\infty} t^r \cdot \frac{e^{-\mu}\mu^r}{r!} = e^{-\mu}\sum_{r=0}^{\infty} t^r \cdot \frac{\mu^r}{r!}$$

$$= e^{-\mu}\sum_{r=0}^{\infty} \frac{(t\mu)^r e^{-t\mu}}{r!} \cdot e^{t\mu}$$

$$= e^{\mu(t-1)}$$

$$\Rightarrow G_{Y_1}(t) = \boxed{e^{\mu(t-1)}}$$

$$M_{Y_1}(t) = E[e^{Y_1 t}] = \sum_{k=0}^{\infty} e^{tk} \frac{\mu^k e^{-\mu}}{k!}$$

$$= e^{-\mu}\sum_{k=0}^{\infty} e^{tk} \frac{\mu^k}{k!}$$

$$= e^{-\mu}\sum_{k=0}^{\infty} \frac{(\mu e^t)^k \cdot e^{-\mu e^t}}{k! \cdot e^{-\mu e^t}}$$

$$= e^{-\mu} \cdot e^{\mu e^t} = \boxed{e^{\mu(e^t-1)}}$$

(ii) $M_S(t) = E[e^{St}] = E[e^{Y_1 t + Y_2 t + \cdots + Y_n t}]$

$= \{E[e^{Y_1 t}]\}^n = e^{n\mu(e^t - 1)}$

Based on 1-1b, S is also a poisson dist$^{\underline{n}}$. with mean $= n\mu$.

(iii)

$G_{X+Y_1}(t) = E[t^{X+Y_1}] = E[E[t^{X+Y_1} | Y_1]]$

$= E[t^{Y_1} E[t^X | Y_1]]$

$= E[t^{Y_1} e^{Y_1(t-1)}] = E[(te^{t-1})^{Y_1}]$

$= \exp[\mu(te^{t-1} - 1)]$

#

Question 9.

(i) The transformation $x = uv$, $y = u - uv$ has Jacobian

$$J = \begin{vmatrix} v & u \\ 1-v & -u \end{vmatrix} = -u$$

$\Rightarrow |J| = |u|$, $\Rightarrow f_{u,v}(u,v) = ue^{-u}$ for $0 \le u < \infty$, $0 \le v \le 1$

$\Rightarrow$ U and V are indep.

& $f_V(v) = 1$ on $[0, 1]$.

(ii) let $w = x$, $z = (y - \rho x)/\sqrt{1-\rho^2}$

$\Rightarrow x = w$, $y = \rho w + z\sqrt{1-\rho^2}$

$$J = \begin{vmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{vmatrix} = \sqrt{1-\rho^2}$$

given the mapping is one - one,

$$\Rightarrow f_{W,Z}(w, z) = \frac{\sqrt{1-\rho^2}}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(1-\rho^2)(w^2+z^2)\right]$$

$$= \frac{1}{2\pi} e^{-\frac{1}{2}(w^2+z^2)} = \frac{1}{2\pi} e^{-\frac{1}{2}w^2} e^{-\frac{1}{2}z^2}$$

$\Rightarrow$ $w$, $z$ are indep $N(0,1)$

(iii) event $\{X>0, Y>0\}$

$= \{w>0,\ z>-w\rho/\sqrt{1-\rho^2}\}$ polar coordinates

$$\Rightarrow P[X>0, Y>0] = \int_{\theta=\alpha}^{\frac{1}{2}\pi} \int_{r=0}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}r^2} r\,dr\,d\theta$$

$$= \int_{\alpha}^{\frac{1}{2}\pi} \frac{1}{2\pi}\,d\theta, \quad \alpha = -\tan^{-1}(\rho/\sqrt{1-\rho^2})$$
$$= -\sin^{-1}\rho$$

$$\Rightarrow P[X>0, Y>0] = \frac{1}{2\pi}\theta \Big|_{\alpha}^{\frac{1}{2}\pi} = \frac{1}{4} + \frac{1}{2\pi}\cdot\sin^{-1}\rho$$

#

Question 10.

(i) $H_0: \mu = 6$, $H_1: \mu \neq 6$

$\bar{x} \sim N(\mu, \frac{\sigma^2}{n}) = N(6, 0.0230)$

$\frac{1}{30} \sum_{i} \sum_{j} y_{ij} = \frac{1}{30} \times 183 = 6.1.$

test statistic $= \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{6.1 - 6}{\sqrt{0.023}}$

$= 0.6593 < 1.96$

$\Rightarrow$ Do not reject the null hypothesis

(ii)

1. Independence of observations: this is an assumption of the model that simplifies the statistical analysis.
2. Normality: the distribution of the residuals are normal
3. Equality of variances: the variance of data in groups should be the same
4. $\sum \tau_i = 0$.

(iii) Sum of square errors between:

$\sum_{i=1}^{3} n_i(\mu_i - \mu)^2 = 15.584.$

sum of square errors within:

$$\sum_{i=1}^{3}(n_i-1)s_i^2 = 8.096.$$

Degree of feedom between: $3-1=2$.

Degree of freedom within: $df_{total} - df_{between}$
$= 29 - 2 = 27$.

Mean square error between:

$$\frac{\text{Sum of square error between}}{2} = \frac{15.584}{2} = 7.792$$

Mean square error within

$$= \frac{\text{Sum of squaree error within}}{27.} = \frac{8.096}{27} = 0.2999$$

$$F \text{ test statistic} = \frac{7.792}{0.2999} = 25.9820. > F_{2,27}$$

$H_0: \mu_{se} = \mu_{ve} = \mu_{vi}$; $H_1: \mu_i \neq \mu_j$ for at least one pair of Se, Ve + Vi.

$\Rightarrow$ there is a difference between at least one pair of sepal lengths.