第一题

$P(X=0)=\frac{1}{3}+\frac{1}{3}=\frac{2}{3}$, $P(X=1)=0+\frac{1}{3}=\frac{1}{3}$
$P(Y=0)=\frac{1}{3}+0=\frac{1}{3}$, $P(Y=1)=\frac{1}{3}+\frac{1}{3}=\frac{2}{3}$

(a) $H(X)=-\sum P(X=x_i)\log_2 P(X=x_i)=\log_2 3-\frac{2}{3}$ $\quad H(Y)=-\sum P(Y=y_i)\log_2 P(Y=y_i)=\log_2 3-\frac{2}{3}$

(b) $H(X|Y)=\sum P(Y=y_i)H(X|Y=y_i)=\frac{2}{3}$
$H(Y|X)=\sum P(X=x_i)H(Y|X=x_i)=\frac{2}{3}$

(c) $H(X,Y)=H(Y|X)+H(X)=\log_2 3$

(d) $I(X,Y)=H(X)-H(X|Y)=\log_2 3-\frac{4}{3}$

第二题

$$0.4\mathcal{N}\left(\begin{bmatrix}10\\2\end{bmatrix},\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)\ +\ 0.6\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}8.4&2.0\\2.0&1.7\end{bmatrix}\right)$$

(a) 从混合的高斯分布中提取出每一维对应的均值和方差得到：
第一维 x_1 的边际分布为 $0.4\mathcal{N}(10,1)+0.6\mathcal{N}(0,8.4)$
第二维 x_2 的边际分布为 $0.4\mathcal{N}(2,1)+0.6\mathcal{N}(0,1.7)$

(b) $\mu=0.4\begin{pmatrix}10\\2\end{pmatrix}+0.6\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}4\\0.8\end{pmatrix}$

第三题

(a) $\because y=Ax+b+\omega$
$\therefore$ 在 x 固定的情况下，$Ax+b$ 为常数
$\because \omega\sim\mathcal{N}(\omega|0,Q)$
$\therefore p(y|x)\sim\mathcal{N}(y|Ax+b,Q)$

(b) $\because x$ 与 ω 相互独立
$\therefore Ax+b$ 与 ω 相互独立

$\therefore \mu_y=E[Ax+b+\omega]=AE[x]+b+E[\omega]=A\mu+b$
$\Sigma_y=Cov[y]=Cov[Ax+b+\omega]=Cov[Ax]+Cov[\omega]=ACov[x]A^T+Q=A\Sigma A^T+Q$

第四题

已知 $H(x)=\int P(x)\log_2 P(x)dx$，且 $\mu=\int xP(x)dx$ 固定，$\sigma^2=\int(x-\mu)^2P(x)dx$ 固定
写为优化问题：

$$\begin{aligned}\max\quad & -\int P(x)\log_2 P(x)dx\\ s.t.\quad & \int P(x)dx=1\\ & \int xP(x)dx=\mu\\ & \int(x-\mu)^2P(x)dx=\sigma^2\end{aligned}$$

构造拉格朗日函数：

$$
\begin{aligned}
L(P(x),\lambda_1,\lambda_2,\lambda_3) = &-\int P(x)\log_2 P(x)dx + \lambda_1\left(\int P(x)dx - 1\right) \\
&+ \lambda_2\left(\int xP(x)dx - \mu\right) + \lambda_3\left(\int (x-\mu)^2 P(x)dx - \sigma^2\right)
\end{aligned}
$$

令 L 对 $P(x)$ 的偏导数为 0（为了计算方便，这里取自然对数 ln），来寻找使 L 最大的 $P(x)$：

$$
\frac{\partial L}{\partial P(x)} = -(\ln P(x) + 1) + \lambda_1 + \lambda_2 x + \lambda_3(x-\mu)^2 = 0
$$

整理得：

$$
P(x) = e^{\lambda_1 - 1 + \lambda_2 x + \lambda_3 (x-\mu)^2}
$$

可以看出 $P(x)$ 的指数部分是关于 x 的二次函数，根据约束条件确定具体参数：

首先，由于均值被固定为 μ，分布必须关于 μ 对称，这意味着 x 的一次项系数 λ_2 必须为 0。此时表达式简化为 $P(x) = e^{\lambda_1 - 1} e^{\lambda_3 (x-\mu)^2}$。

其次，利用方差约束 $\int (x-\mu)^2 P(x)dx = \sigma^2$，可解得 $\lambda_3 = -\frac{1}{2\sigma^2}$。

最后，利用归一化约束 $\int P(x)dx = 1$，可解得常数项 $e^{\lambda_1 - 1} = \frac{1}{\sqrt{2\pi}\sigma}$。

将参数代入，最终得到：

$$
P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
$$

即高斯分布是该约束下的最大熵分布。

第五题

设从哪个袋子中挑选水果为随机变量 X（X 可能的取值为 1, 2）。

设挑选到的水果为随机变量 Y（1 代表芒果，2 代表苹果），则：P(X=1)=0.6,　P(X=2)=0.4

由贝叶斯公式：

$$
\begin{aligned}
P(X=2 \mid Y=1) &= \frac{P(Y=1 \mid X=2)P(X=2)}{P(Y=1)} \\
&= \frac{P(Y=1 \mid X=2)P(X=2)}{P(Y=1 \mid X=1)P(X=1) + P(Y=1 \mid X=2)P(X=2)} \\
&= \frac{\dfrac{1}{2}\cdot 0.6}{\dfrac{1}{2}\cdot 0.6 + \dfrac{2}{3}\cdot 0.4} \\
&= \frac{1}{3}
\end{aligned}
$$