1. P(at least 1 heads) $= 1 -$ P(no heads) $= 1 - \frac{1}{2^{10}}$.

2. $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.

3. $\mathrm{P}(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - P(A_i \cup_{j=1}^{m} A_j)$
   $1 \geq P(A_1 \cap A_j) \geq 0 forall A_i and A_j$
   => because the $\mathrm{P}(A_i \cap A_j)$will always be positive, subtracting it will always reduce the probability or do nothing to it.
   It can never grow.
   $=> P(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$
   the equality holds if all events are mutually exclusive from every other event

4. (a) i. $\mathrm{P}(x = 0) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10}$

   $\mathrm{P}(x = 1) = \frac{3}{10} + \frac{4}{10} = \frac{7}{10}$

   $\mathrm{P}(y = 0) = \frac{1}{10} + \frac{3}{10} = \frac{4}{10}$

   $\mathrm{P}(y = 1) = \frac{2}{10} + \frac{4}{10} = \frac{6}{10}$

   ii. $\mathrm{P}(x = 0|y = 0) = \frac{1}{10} * \frac{10}{4} = \frac{10}{40} = \frac{1}{4}$

   $\mathrm{P}(x = 1|y = 0) = \frac{3}{10} * \frac{10}{4} = \frac{30}{40} = \frac{3}{4}$

   $\mathrm{P}(x = 0|y = 1) = \frac{1}{10} * \frac{10}{6} = \frac{10}{60} = \frac{1}{6}$

   $\mathrm{P}(x = 1|y = 1) = \frac{3}{10} * \frac{10}{6} = \frac{30}{60} = \frac{1}{2}$

   $\mathrm{P}(y = 0|x = 0) = \frac{1}{10} * \frac{10}{3} = \frac{10}{30} = \frac{1}{3}$

   $\mathrm{P}(y = 1|x = 0) = \frac{2}{10} * \frac{10}{3} = \frac{20}{30} = \frac{2}{3}$

   $\mathrm{P}(y = 0|x = 1) = \frac{3}{10} * \frac{10}{7} = \frac{30}{70} = \frac{3}{7}$

   $\mathrm{P}(y = 1|x = 1) = \frac{4}{10} * \frac{10}{7} = \frac{40}{70} = \frac{4}{7}$

   iii. $\mathbb{E}(x) = 0 * P(x = 0) + 1 * P(x = 1) = \frac{7}{10}$

   $\mathbb{E}(Y) = 0 * P(y = 0) + 1 * P(y = 1) = \frac{6}{10} = \frac{3}{5}$

   $\mathbb{V}(x) = \mathbb{E}(X^2) - \mathbb{E}(x)^2, \mathbb{E}(x^2) = 0^2 * P(x = 0) + 1^2 * P(x = 1) = \frac{7}{10}$

   $=> V(x) = \frac{7}{10} - \frac{7}{10}^2 = \frac{70}{100} - \frac{49}{100} = \frac{21}{100}$

   $\mathbb{E}(Y^2) = 0^2 * P(y = 0) + 1^2 * P(y = 1) = \frac{3}{5}$

   $=> V(Y) = \frac{3}{5} - \frac{3}{5}^2 = \frac{15}{25} - \frac{9}{25} = \frac{6}{25}$

   iv. $\mathbb{E}(Y|x = 0) = 0 * P(y = 0|x = 0) + 1 * P(y = 1|x = 0) = \frac{2}{3}$

$\mathbb{E}(Y|x=1) = 0 * P(y=0|x=1) + 1 * P(y=1|x=0) = \frac{4}{7}$

$\mathbb{V}(Y|x=0) = \mathbb{E}((Y|x=0)^2) - \mathbb{E}(Y|x=0)^2$

$\mathbb{E}((Y|x=0)^2) = 1^2 * P(y=1|x=0) = \frac{2}{3}$

$=> \mathbb{V}(Y|x=0) = \frac{2}{3} - (\frac{2}{3})^2 = \frac{6}{9} - \frac{4}{9} = \frac{2}{9}$

$\mathbb{E}((Y|x=1)^2) = 1^2 * P(y=1|x=0) = \frac{4}{7}$

$=> \mathbb{V}(Y|x=1) = \frac{4}{7} - (\frac{4}{7})^2 = \frac{28}{49} - \frac{16}{49} = \frac{12}{49}$

v. $\text{COV}(XY) = \mathbb{E}(xY) - \mathbb{E}(x)\mathbb{E}(Y)$

$\mathbb{E}(xY) = \frac{4}{10}$

$=> \text{COV}(xY) = \frac{4}{10} - (\frac{7}{10})(\frac{3}{5}) = \frac{20}{50} - \frac{21}{50} = \frac{-1}{50}$

(b) because $\text{COV}(xY) \neq 0$, they are not independent

(c) when x is not assigned a specific value, $\mathbb{E}(Y|x)\&\mathbb{V}(Y|x)$are not constant because Y is not independent from x.

5. $X\ \mathcal{N}(x|0,1)$

(a) $\mathbb{E}(Y) = \mathbb{E}(e^x)$,
$M_x(t) = \mathbb{E}(e^{tx}) = e^{0t+\frac{1}{2}*1^2*t^2} = e^{\frac{t^2}{2}}$
$\mathbb{E}(Y) = M_x(1) = e^{\frac{1}{2}}$

(b) $\mathbb{E}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2$
$Y^2 = e^{2x}$
$\mathbb{E}(Y^2) = M_x(2) = e^{\frac{2^2}{2}} = e^2$
$V(Y) = e^2 - (e^{\frac{1}{2}})^2 = e^2 - e$

6. (a) $\mathbb{E}(\mathbb{E}(Y|X) = \mathbb{E}(Y)$

$\mathbb{E}(Y) = \int_{-\infty}^{\infty} yf(y)\,dy$

$\mathbb{E}(Y|X=x) = \int_{-\infty}^{\infty} yf_{y|x}(Y|X)\,dy$

$\mathbb{E}(\mathbb{E}(Y|X)) = \int_{-\infty}^{\infty}(\int_{-\infty}^{\infty} yf_{Y|X}(Y|X)dy)\,dy$

$= \int_{-\infty}^{\infty}(\int_{-\infty}^{\infty} yf_{X,Y}(X,Y)dx)\,dy$

$= \int_{-\infty}^{\infty} yf_Y(Y)dy\ = \mathbb{E}(Y)$

$=> \mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$

(b) $\mathbb{V}(Y) = \mathbb{E}(\mathbb{V}(Y|X))$

$\mathbb{V}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2$

$\mathbb{E}(\mathbb{V}(Y|X)) = \mathbb{E}(\mathbb{E}(Y^2|X) - \mathbb{E}(Y|X)^2)$
$= \mathbb{E}(\mathbb{E}(Y^2|X)) - \mathbb{E}(\mathbb{E}(Y|X)^2)$
$= \mathbb{E}(Y^2) - \mathbb{E}(\mathbb{E}(Y|X)^2)$

$\mathbb{V}(\mathbb{E}(Y|X)) = \mathbb{E}(\mathbb{E}(Y|X)^2) - \mathbb{E}(\mathbb{E}(X|Y))^2$
$= \mathbb{E}(\mathbb{E}(Y|X)^2) - \mathbb{E}(Y)^2$

$\mathbb{V}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(\mathbb{E}(Y|X)^2) + \mathbb{E}(\mathbb{E}(Y|X)^2) - \mathbb{E}(Y)^2 = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2$

7. (a) $f(x) = \frac{1}{1+e^{-a^Tx}} => f(z) = \frac{1}{1+e^{-z}}$

$\nabla f(x) = \frac{df(x)}{dx} = \frac{df(z)}{dz}\nabla z$

$\frac{df(z)}{dz} = \frac{d}{dz}(\frac{1}{g(z)}, V(z) = 1 + e^{-z}, U(z) = 1$

$\frac{dU(z)}{dz} = 0, \frac{dV(z)}{dz} = -e^{-z}$

$\frac{df(z)}{dz} = \frac{(1+e^{-z})*0-(1*-e^{-z})}{(1+e^{-z})^2} = \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1}{1+e^-z} * \frac{e^{-z}}{1+e^{-z}} = f(z) * (1 - f(z))$

$\nabla z = \frac{dz}{dx} = a$

$=> \nabla f(x) = f(x)(1 - f(x)) * a$

(b) $\nabla f(x) = f(x)(1 - f(x)) * a$

$\nabla^2 f(x) = \nabla(\nabla f(x)) = \frac{d}{dx}(f(x)(1 - f(x) * a)$

$(\frac{df(x)}{dx})(1 - f(x))(a) + (f(x)(\frac{-df(x)}{dx})(a) + f(x)(1 - f(x))(0), \frac{df(x)}{dx} = f(x)(1 - f(x)) * a$

$- > (f(X)(1 - f(x))a)(1 - f(x))(a) + f(x)(-(f(x)(1 - f(x))(a)))(a)$

$=> \nabla^2 f(x) = a^2(f(x) - 3f(x)^2 + 2f(x)^3)$

(c) $a = [1, 1, 1, 1, 1]^T, x = [0, 0, 0, 0, 0]^T$

$\nabla f(x) = f(x)(1 - f(x))a$

$e^{[1,1,1,1,1]^T*[0,0,0,0,0]} = e^0 = 1$

$f(x) = \frac{1}{1+e^0} = \frac{1}{2}$

$=> \nabla f([0, 0, 0, 0, 0) = \frac{1}{2}(1 - \frac{1}{2} * [1, 1, 1, 1, 1]^T = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}]^T$

(d) $\nabla^2 f(x) = [1, 1, 1, 1, 1]^{T2}(\frac{1}{2} - 3(\frac{1}{2})^2 + 2(\frac{1}{2})^3)$

$= 1 * (\frac{4}{8} - \frac{6}{8} + \frac{2}{8}) = 1 * 0 = 0$

8. $g(x) = -log(f(x)) = log(1 + e^{-a^T x})$

$\nabla g(x) = \frac{d}{dx} log(u(x)) = \frac{1}{u(x)} * \nabla u(x)$

$\nabla u(x) = -e^{-a^T x} * -a$

$\nabla g(x) = \frac{1}{1+e^{-a^T x}} * (-e^{-a^T x} * -a) = \frac{(e^{-a^T x} * a)}{1+e^{-a^T x}}$

$= (1 - f(x)) * a$

$\nabla^2 g(x) = \nabla(1 - f(x)) * a = (\frac{-df(x)}{dx} * a + (1 - f(x)) * 0$

$=> \nabla^2 g(x) = f(x) * (1 - f(x)) * a^2$

g(x) will always be convex because $\nabla^2 g(x)$will always be greater than 0 for all $0 \leq f(x) \leq 1$