1. First note that $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$ implies that $S = \sum X_i \sim \text{Poisson}(n\lambda)$. (This is most easily shown using moment generating functions.)

In order to show that $S$ is sufficient by the definition, we must show that the conditional distribution

$$f_{X_1, X_2, \ldots, X_n | S}(x_1, x_2, \ldots, x_n | s)$$

does not depend on $\lambda$. Since the random variables involved here are discrete, this is the same thing as

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n | S = s)$$

which is

$$\frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, S = s)}{P(S = s)}$$

(1)

Now since $S = \sum X_i$, if the fixed value $s$ isn’t equal to the sum $\sum x_i$, the numerator of (1) (and therefore all of (1)) is zero.

Now assume that $s = \sum x_i$. Then

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, S = s) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

much in the same way that $P(X_1 = 2, X_2 = 5, X_1 + X_2 = 7) = P(X_1 = 2, X_2 = 5)$, for example.

Therefore,

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n | S = s) = \frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n, S = s)}{P(S = s)}$$

$$= \frac{P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)}{P(S = s)}$$

$$\overset{indep}{=} \frac{P(X_1 = x_1) \cdot P(X_2 = x_2) \cdots P(X_n = x_n)}{P(S = s)}$$

$$= \frac{e^{-\lambda x_1} \cdots e^{-\lambda x_n}}{x_1! \cdots x_n!} \cdot \frac{e^{-n\lambda S} \sum x_i}{e^{-n\lambda (n\lambda) \sum x_i}} \cdot \frac{\prod (x_i !)}{\prod \sum x_i !} \cdot \frac{(\sum x_i) !}{n \sum x_i \prod (x_i !)}.$$ 

In either case, ($s \neq \sum x_i$ or $s = \sum x_i$), the conditional density for $X_1, X_2, \ldots, X_n$ given $S$ does not depend on the parameter $\lambda$. Therefore $S = \sum X_i$ is sufficient for the Poisson model.
2. (a)

\[ f(x; \mu) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2}(x-\mu)^2} \]

\[ \Rightarrow f(\vec{x}; \mu) \overset{\text{iid}}{=} \prod_{i=1}^{n} f(x_i; \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \]

\[ = (2\pi \sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum (x_i-\mu)^2} \]

Now

\[ \sum (x_i - \mu)^2 = \sum x_i^2 - 2\mu \sum x_i + n\mu^2, \]

so breaking the joint pdf out into a “data part” and a “data mixed with parameter part” (keep in mind that \( \sigma^2 \) is fixed and known and not being thought of as a parameter) gives

\[ f(\vec{x}; \mu) = (2\pi \sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum x_i^2} \cdot e^{-\frac{1}{2\sigma^2}(-2\mu \sum x_i+n\mu^2)}. \]

Everything before the “multiplication dot” can be thought of as the \( h(\vec{x}) \) part and everything after can be thought of as the \( g(s(\vec{x}); \mu) \) part. Is is this later part we look at in order to determine a sufficient statistic. Since the \( x_i \)’s appear in \( g \) through the statistic \( \sum x_i \), we have that

\[ S = \sum X_i \]

is a sufficient statistic for \( \mu \).

So, when estimating \( \mu \), it is sufficient or “enough” to have \( \sum X_i \) instead of the entire data set \( X_1, X_2, \ldots, X_n \). It is also enough (sufficient) to have \( \bar{X} \) since you can recover the value of \( \sum X_i \) from the value of \( \bar{X} \). Another way to think about it is you can write \( \sum X_i = n\bar{X} \) so

\[ f(\bar{x}; \mu) = (2\pi \sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum x_i^2} \cdot e^{-\frac{1}{2\sigma^2}(-2\mu n\bar{X}+n\mu^2)}. \]

and you can cite the factorization criterion directly.

(b) Now break up the joint pdf \( f(\vec{x}; \sigma^2) \) this way:

\[ f(\vec{X}, \sigma^2) = (2\pi)^{-n/2} \cdot (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum (x_i-\mu)^2}. \]

Everything before the “multiplication dot” can be thought of as the \( h(\vec{x}) \) part and everything after can be thought of as the \( g(s(\vec{x}); \mu) \) part. (Although we may put that \( (2\pi)^{-n/2} \) part in with the \( g \) function and say that \( h(\vec{x}) = 1. \) Is is this later part we look at in order to determine a sufficient statistic. Since the \( x_i \)’s appear in \( g \) through the statistic \( \sum (x_i - \mu)^2 \), we have that

\[ S = \sum (X_i - \mu)^2 \]

is a sufficient statistic for \( \sigma^2 \).
(c) Now break up the joint pdf $f(\vec{x}; \mu, \sigma^2)$ this way:

$$f(\vec{X}, \sigma^2) = (2\pi)^{-n/2} \cdot (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum(x_i - \mu)^2}.$$  

and multiply it out

$$f(\vec{X}, \sigma^2) = (2\pi)^{-n/2} \cdot (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum x_i^2 + \frac{\mu}{\sigma^2} \sum x_i - \frac{n\mu^2}{2\sigma^2}}.$$  

Now take $h(\vec{x})$ to be $(2\pi)^{-n/2}$ (or identically 1), and take $g((s\vec{x}); \mu, \sigma^2)$ to be the rest. Here $s$ is actually vector-valued, i.e: we have $g(s(\vec{x}); \mu, \sigma^2)$.

The data appears in $g$ in “two chunks”: $\sum x_i^2$ and $\sum x_i$. So,  

$$\vec{S} = \vec{s}(\vec{X}) = (\sum X_i^2, \sum X_i)$$

is a vector of jointly sufficient statistics for $\mu$ and $\sigma^2$.

Now $\vec{X}$ and $S^2$ are also jointly sufficient for $\mu$ and $\sigma^2$ since from $\vec{X}$, you can recover the value of $\sum X_i$ and then using this $\sum X_i$ and $S^2$ you can recover $\sum X_i^2$. So, it is enough to know $\vec{X}$ and $S^2$.

3. (a) 

$$f(x; p) = p^x (1 - p)^{1-x} \cdot I_{\{0,1\}}(x)$$

$$\Rightarrow f(\vec{x}; p) = p^{\sum x_i} (1 - p)^{n - \sum x_i} \cdot \prod I_{\{0,1\}}(x_i)$$

Take the $h$ function to be the product of the indicators (or to be 1 in the case that you did not include the indicators). Take $g$ to be the rest of it where you have parameter mixed with data. The data appears in $g$ through the sum $\sum x_i$. Therefore,

$$S = \sum X_i$$

is sufficient for $p$.

(b) 

$$E[I_{X_1=1}] = P(X_1 = 1) = p \quad \text{unbiased} \quad \checkmark$$

(c) We want $E[I_{X_1=1}|S]$. We start by fixing $S$ to be some value, say "$a$", computing $E[I_{X_1=1}|S = s]$, and then plugging the random (capital) $S$ back in for $s$.

$$E[I_{X_1=1}|S = s] = P(X_1 = 1|S = s)$$

$$= P(X_1 = 1|X_1 + Y = s)$$
where $Y = \sum_{i=2}^{n}$. Note that $X_1$ is independent of $Y$ and that $Y \sim bin(n-1, p)$. Well,

$$P(X_1 = 1|X_1 + Y = s) = \frac{P(X_1=1,X_1+Y=s)}{P(X_1+Y=s)}$$

$$= \frac{P(X_1=1,Y=s-1)}{P(\sum_{i=1}^{n} X_i=s)}$$

$$\overset{iid}{=} \frac{P(X_1=1)\cdot P(Y=s-1)}{P(\sum_{i=1}^{n} X_i=s)}$$

Since $X_1 \sim Bernoulli(p)$, $Y \sim bin(n-1, p)$ and $\sum_{i=1}^{n} X_i \sim bin(n, p)$, we have

$$p \cdot \left( \frac{n-1}{s} \right) / \left( p^{s-1}(1-p)^{(n-1)-(s-1)} \right) \cdot \frac{s}{n} = \frac{s}{n}$$

Since

$$E[I_{X_1=1}|S = s] = \frac{s}{n},$$

we have that

$$E[I_{X_1=1}|S] = \frac{S}{n} = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}.$$  

(d) 

$$E[E[I_{X=1}=1]] = E[\overline{X}] = p$$

since $X_1, X_2, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$.

(e) 

$$Var[\hat{p}_1] = Var[\overline{I}_{X_1=1}]$$

$$= E[I_{\{X_1=1\}}^2] - \left( E[I_{\{X_1=1\}}] \right)^2$$

$$= p - p^2 = p(1-p)$$

$$Var[\hat{p}_2] = Var[\overline{X}] = \frac{Var(X_1)}{n} = \frac{p(1-p)}{n}$$

So, we see that the Rao-Blackwell Theorem gave us something that is still unbiased but with smaller variance!

4. This problem has been removed from the assignment. (It is an inverse Fourier transform problem which is “beyond the scope” of the course.)
5. The pdf is 
\[ f(x; \theta) = \theta x^{\theta-1} I_{(0,1)}(x). \]

The joint pdf is 
\[ f(\vec{x}; \theta) = \theta^n \prod_{i=1}^{n} x_i^{\theta-1} \prod_{i=1}^{n} I_{(0,1)}(x_i) \]
\[ = \theta^n [\prod_{i=1}^{n} x_i]^{\theta-1} \prod_{i=1}^{n} I_{(0,1)}(x_i) \]
\[ = \frac{\theta^n}{\exp[\theta]} \prod_{i=1}^{n} I_{(0,1)}(x_i) \exp[\sum_{i=1}^{n} \ln x_i] \]

So, by “one-parameter exponential family”,
\[ S = \sum_{i=1}^{n} \ln X_i \]
is complete and sufficient for \( \theta \).

(a) We need to find a function of \( S \) that is unbiased for \( 1/\theta \). Letting \( y = \ln x = g(x) \), a simple “\( g \)-inverse” transformation shows us that 
\[ f_Y(y) = \theta e^{\theta y} I_{(-\infty,0)}(x). \]

This is similar to an exponential distribution. To make it an actual exponential distribution (easier to work with), we will instead take \( Y = -\ln X \). Now \( Y \sim exp(rate = \theta) \). Therefore,
\[ S = -\sum_{i=1}^{n} \ln X_i = -W \]
where \( W \sim \Gamma(n, \theta) \).

So,
\[ E[S] = -E[W] = -\frac{n}{\theta} \]
which implies that
\[ \hat{\theta} = -\frac{1}{n} S = -\frac{1}{n} \ln X_i \]
is the UMVUE for \( \theta \).
(b) We now need to find a function of $S$ whose expected value is $\tau(\theta) = (\theta/(\theta + 1))^n$. Note that this looks like the moment generating function of a $\Gamma(n, \theta)$, evaluated at $t = -1$. Indeed,

$$E[e^S] = E[e^{-W}] = M_W(-1) = \left(\frac{\theta}{\theta + 1}\right)^n.$$ 

Thus, 

$$\widehat{\tau}(\theta) = e^S = e^{\sum \ln X_i} = \prod X_i$$

is the UMVUE for $\tau(\theta)$.

6.  

$$f(\vec{x}; \gamma) = \prod_{i=1}^n \frac{\gamma}{(1 + x_i)^{\gamma+1}} I_{(0, \infty)}(x_i)$$

$$= \gamma^n \cdot \prod_{i=1}^n I_{(0, \infty)}(x_i) \cdot \prod_{i=1}^n \left[(1 + x_i)^{-(\gamma+1)}\right]$$

$$= \gamma^n \cdot \prod_{i=1}^n I_{(0, \infty)}(x_i) \cdot \exp \left[-(\gamma + 1) \sum_{i=1}^n \ln(1 + x_i)\right]$$

So, by “one-parameter exponential family” (with $a(\gamma) = \gamma^n$, $b(\vec{x}) = \prod_{i=1}^n I_{(0, \infty)}(x_i)$, $c(\gamma) = - (\gamma + 1)$, and $d(\vec{x}) = \sum_{i=1}^n \ln(1 + x_i)$), we have that 

$$S = d(\vec{X}) = \sum_{i=1}^n \ln(1 + X_i)$$

is complete and sufficient for $\gamma$.

(a) We need to find a function of $S$ whose expected value is $1/\gamma$. Let’s begin by considering $S$ itself.

The simple “g-inverse” transformation technique gives us that 

$$Y_i = \ln(1 + X_i) \sim exp(rate = \gamma).$$

Hence 

$$S = \sum_{i=1}^n \ln(1 + X_i) \sim \Gamma(1, 1/\gamma).$$

So,

$$E[S] = n \cdot \frac{1}{\gamma}.$$ 

Therefore, the UMVUE for $\tau(\gamma) = 1/\gamma$ is 

$$\widehat{\tau}(\gamma) = \frac{1}{n} S = \frac{1}{n} \frac{1}{\sum_{i=1}^n \ln(1 + X_i)}.$$
(b) Now we need to find a function of $S$ whose expected value is $\gamma$. We know that $E[1/S] \neq 1/E[S]$, but, motivated by wanting to “flip the $\gamma$”, we are going to try $E[1/S]$ anyway:

\[
E \left[ \frac{1}{S} \right] = \int f_{S}(s) \, ds
\]

\[
= \int_{0}^{\infty} \frac{1}{s} \frac{1}{\Gamma(n)} \gamma^{n} s^{n-1} e^{-\gamma s} \, ds
\]

\[
= \int_{0}^{\infty} \frac{1}{\Gamma(n)} \gamma^{n} s^{n-2} e^{-\gamma s} \, ds
\]

\[
= \frac{1}{\Gamma(n)} \Gamma(n-1) \gamma \int_{0}^{\infty} \frac{1}{\Gamma(n-1)} \gamma^{n-2} s^{n-2} e^{-\gamma s} \, ds
\]

\[
= \frac{1}{\Gamma(n)} \Gamma(n-1) \gamma \cdot 1
\]

\[
= \frac{1}{n-1} \gamma
\]

So

\[
\hat{\gamma}_{UMVUE} \left( \frac{n-1}{S} \right) = \frac{n-1}{\sum_{i=1}^{n} \ln(1 + X_{i})}.
\]

7. First note that, since $S$ is sufficient and $M$ is minimal sufficient, by definition, $M$ must be a function of $S$.

(a) We know that $E[S|M]$ is a function of $M$, but since $M$ is a function of $S$, we now have that $E[S|M]$ is a function of $S$. Therefore, $S - E[S|M]$ is a function of $S$.

(b) $E[S - E[S|M]] = E[S] - E[E[S|M]] = E[S] - E[S] = 0$

(c) Since $S$ is complete, part (b) implies that $S - E[S|M] = 0$, or $S = E[S|M]$. (By definition of completeness with $g(S) = E[S|M]$.)

(d) By part (c), $S = E[S|M]$ is a function of $M$ since $E[S|M]$ is a function of $M$.

(e) $M$ minimal sufficient implies that $M$ is sufficient and can be written as a function of every other set of sufficient statistics. We have just shown that $S$ is a function of $M$, hence $S$ is a function of a function of every other set of sufficient statistics and is therefore can be written as a function of every other set of sufficient statistics. Hence, $S$ is minimal sufficient!