1. The following values:

0.312, 0.238, 0.446, 0.968, 0.576, 0.471, 0.596

were generated using R by the command

```
x <- rgamma(7, alpha, lambda)
```
to give 7 i.i.d. variables with a gamma($\alpha, \lambda$) distribution for some numerical values of $\alpha$ and $\lambda$ that I gave. Estimate what those values were by method of moments.

2. Suppose in two independent trials with probability $p$ of success on each we observe $X$ successes.

(a) Find an unbiased estimator $T(X)$ of the function $g(p) = p^2$. \textit{Hint:} $X$ has just three possible values 0, 1, 2, so the estimator is given by the three numbers $T(0)$, $T(1)$, and $T(2)$. For each $p$, each value of $X$ has a certain binomial probability. So the condition for $T(X)$ to be unbiased gives an equation that has to be satisfied for all $p$ with $0 < p < 1$.

(b) Does the equation in the hint give unique solutions for $T(j)$, $j = 0, 1, 2$, and what solution(s) do you find?

(c) What is most surprising about the results of part (b)?

3. A beta($a, b$) distribution has a density

$$f_{a,b}(x) = x^{a-1}(1-x)^{b-1}/B(a,b)$$

for $0 < x < 1$ and 0 elsewhere, for $0 < a < +\infty$ and $0 < b < +\infty$, and $B(a,b)$ is the beta function, which normalizes the density to be a probability density and satisfies $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ where $\Gamma$ is the gamma function. If $X$ has this density then $EX = a/(a+b)$. The simplest example of such a distribution is the uniform $U[0,1]$ distribution with $a = b = 1$.

If $p$ is the probability of success in each of $n$ independent trials, suppose $p$ has a prior distribution beta($a, b$). Then if we observe $X$ successes in the $n$ trials, the posterior distribution will be beta($a + X, b + n - X$). The Bayes estimate $T(X, n)$ of $p$, minimizing the risk, i.e. minimizing the integral from 0 to 1 of $(T(X, n) - p)^2$ with respect to the posterior distribution, will just be the integral of $p$ times the posterior density.

(a) For the beta($1/2, 1/2$) prior, what is the Bayes estimator of $p$ as a function of $n$ and $X$?

(b) Compare the estimator from part (a) to the classical estimator $\hat{p} = X/n$ in terms of their squared-error losses $E_p((\hat{p} - p)^2)$ and $E_p((T(X, n) - p)^2)$. For which values of $p$ does each estimator perform better in the sense of having smaller expected loss?

4. Let $X_1, \ldots, X_n$ be i.i.d. having a geometric distribution with for some $p$ such that $0 < p \leq 1$, namely $P(X_1 = k) = (1 - p)^{k-1}p$ for $k = 1, 2, \ldots$.

(a) What is the maximum likelihood estimate (MLE) of $p$ based on $X_1, \ldots, X_n$?

(b) What is the method-of-moments estimate of $p$?

(c) Suppose we view the situation as follows. We have done $S_n = X_1 + \cdots + X_n$ independent trials with probability $p$ of success on each and observed exactly $n$ successes. Then what
is the binomial MLE of $p$? (It isn’t obvious that this should be equivalent to (a), as we did a random number of trials and stopped after the $n$th success.)

5. Consider the family of mixtures of two normal distributions, having densities of the form

$$f(x, \theta) = \frac{\lambda}{\sqrt{2\pi\sigma_1}} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) + \frac{1 - \lambda}{\sqrt{2\pi\sigma_2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right)$$

where $\theta = (\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2)$ is a 5-dimensional parameter with $\mu_1$ and $\mu_2$ any real numbers, $0 < \sigma_j < \infty$ for $j = 1, 2$, and $0 < \lambda \leq 1/2$. Suppose given $n$ observations $X_1, \ldots, X_n$, not all equal, assumed to be i.i.d. from such a distribution. If a value $\theta'$ of a parameter is such that as $\theta$ approaches $\theta'$ (possibly under some restrictions), the likelihood approaches $+\infty$, then we may consider $\theta'$ as a maximum likelihood estimate (MLE) of $\theta$, or the MLE if it’s unique.

(a) For the given family of normal mixture densities, do there exist such $\theta'$? Are they unique? *Hint:* the exponential of a nonpositive number is at most 1, so the likelihood can only approach $+\infty$ if at least one of the $\sigma_j$ approaches 0. But if say $\sigma_1$ approaches 0, then $\exp(-(X_j - \mu_1)^2/(2\sigma_1^2))$ will approach 0 very fast if $\mu_1$ is fixed and unequal to $X_j$. In the likelihood function the $X_j$ are fixed and the parameters are free to vary, so for what value(s) of $\mu_1$ would we get large likelihood as $\sigma_1 \downarrow 0$?

(b) Suppose the observations are really i.i.d. with a density of the given form having $0 < \lambda \leq 1/2$, $0 < \sigma_1 < \infty$, $0 < \sigma_2 < \infty$. How successful will choosing parameters for which the likelihood is very large, as in part (a), be in approximating the actual parameters, supposing we can take $n$ as large as we want? Specifically, supposing $\lambda = 1/2$ is fixed, how well will the distribution function of the distribution with estimated parameters approximate the one for the true parameters?