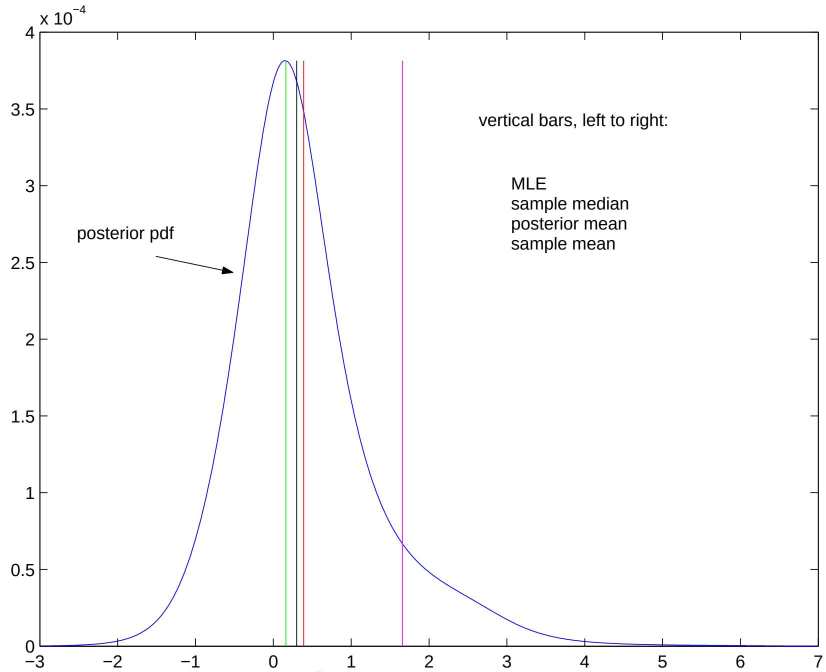
ANSWERS FOR ESTIMATION, TESTING, CONFIDENCE INTERVAL EXERCISE

(1) (a)

(b) The requested plot is shown below, followed by code that constructed it.




```
x=[6.3, -1.1, 0.3, 0, 2.8]
m=-3:.02:7;
pc=prod(1+(ones(5,1)*m-x'*ones(1,size(m,2))).^2);
pc=1../(pi*pc);
plot(m,pc)
hold
max(pc)
top=ans;
pc=pc/sum(pc);
postMean=sum(pc.*m)
plot([postMean,postMean],[0,top],'r')
[pcx,iml]=max(pc)
mle=m(iml)
plot([mle,mle],[0,top],'g')
mean(x)
plot([ans,ans],[0,top],'m')
x
plot([.3,.3],[0,top],'k')
```

(c) The flat-prior posterior pdf is the likelihood, which is

$$\frac{1}{\pi^5 \prod_{i=1}^5 \big(1+(x_i-\mu)^2\big)} .$$

As a function of μ, this expression is one over a polynomial in μ of 6th order. Since the polynomial has no zeros for real μ, it is itself integrable, and for any polynomial $q(\mu)$ of order $p \leq 4$, q times this likelihood function will be integrable, since the product expression will be $O(\mu^{-2})$. This means that integer moments up to 4th order of this posterior pdf will exist.

(d) The sample mean has variance $(1/N)\,\mathrm{Var}(x)$. Since here each individual observation on x has infinite variance (because $x^2/(1+x^2)$ is $O(1)$ as $|x| \to \infty$), the sample mean also has infinite variance, and thus unbounded quadratic loss. Any finite variance estimator will dominate it, therefore. A cheap example of such an estimator is the trivial estimator $\hat{\mu} = 0$. Though this has MSE that grows arbitrarily large as the true value of μ moves away from zero, the MSE is finite, conditional on any specific value of μ.

The fact that the posterior pdf has finite variance, as shown in the first part of the problem, does not imply that the posterior mean has finite sampling variance. A Bayesian posterior mean formed from a proper prior that has non-zero pdf over the whole real line must be admissible, except possibly for sets of Lebesgue measure zero. Therefore, since an estimator with finite MSE exists, posterior mean from a proper prior will have finite MSE and dominates the sample mean. The flat-prior posterior mean is not a Bayes estimator for any proper prior, so we can't use this argument directly to show that it must have finite variance. However, the proper-prior posterior means do form a large class of estimators that must dominate the sample mean, and probably the flat-prior posterior mean (and the sample median, the posterior median, and the MLE, for that matter) do also.

(2) (a) The $\Gamma(p,\alpha)$ distribution has mean p/α and variance p/α^2. It also has finite moments of all orders, so a SLLN will tell us that $\bar{t}$ converges in probability to p/α and the sample second moment (the average of the t_i^2's) converges in probability to $(p^2+p)/\alpha^2$. Then we can apply the theorem given in the hint to conclude that the sample variance s_t^2 converges to p/α^2. Finally we apply the theorem again to conclude that $\bar{t}/s_t^2 \xrightarrow{P} \alpha$ and $\bar{t}^2/\sigma_t^2 \xrightarrow{P} p$.

(b) It is easy to check that the likelihood in this model depends on the data only through $\prod_i t_i$ and $\sum_i t_i$ (or, equivalently, on the sample arithmetic and geometric means), so these are sufficient statistics. The proposed estimators are not functions of these sufficient statistics (because they also depend on the sample second moment). So we should expect that the proposed estimator can be improved upon. Any estimator that depends only on the sufficient statistics will be better in the sense that it cannot be uniformly improved upon by reducing it to a function of sufficient statistics. If the loss function is strictly convex, replacing the proposed estimator by its conditional expectation, given the sufficient statistics, will improve the estimator. However, I don't think an analytic expression for such a conditional expectation is available. It would be a challenging computation even numerically. For a given prior, the posterior means of α and p, which are optimal estimates under quadratic loss, could be calculated by straightforward numerical integration. The presence of the gamma function in the likelihood makes analytic integration of it difficult or impossible.

(3) (a) Denote the proposed test statistic by $w = (t_1 - t_2)/(t_1 + t_2)$. Let $u = t_1 + t_2$. Making the change of variables from t_1, t_2 to u, w gives us the Jacobian

$$\frac{\partial(u,w)}{\partial(t_t,t_2)} = \left| \begin{bmatrix} 1 & 1 \\ \dfrac{1-w}{u} & \dfrac{-1-w}{u} \end{bmatrix} \right| = \frac{-2}{u}\,.$$

The joint pdf of t_1 and t_2 under the equal-α null hypothesis is simply

$$\alpha^2 e^{-\alpha(t_1+t_2)}\, dt_1\, dt_2\,, \quad t_1 > 0, t_2 > 0\,.$$

Making the change of variables to u and w, taking account of the Jacobian, gives us

$$\frac{1}{2} u e^{-u}\, du\, dw\,, \quad u > 0, w \in (-1,1)\,.$$

Because the joint pdf doesn't depend on w, it is clear that w's marginal distribution is uniform over its range of $(-1,1)$. A test of size α can therefore be constructed by letting the rejection region be any interval, or collection of intervals, within $(-1,1)$ whose lengths add up to 1.9. (1.9, not .95, because the density on this interval of length two is .5.) Since we would like to have power against the alternative that the two distributions have different α's, it is clear that what makes sense is to choose the rejection region to be $(-1,-.95) \cup (.95,1)$.

(b) The likelihood times prior pdf within the equal-α parameter space is

$$.5\gamma\alpha^2 e^{-\alpha(t_1+t_2+\gamma)}\,.$$

We can integrate the α out of this easily, because as a function of α it has the form of a $\Gamma(3)$ pdf. Integrating out the α to get the posterior weight on this part of the parameter space gives us

$$\frac{\gamma\Gamma(3)}{(\gamma + t_1 + t_2)^3}\,.$$

Within the unequal-α space the likelihood times prior is

$$.5\gamma^2 \alpha_1 \alpha_2 e^{-\alpha_1(t_1+\gamma)-\alpha_2(t_2+\gamma)}\,.$$

Integrating the two α's out of this expression (again recognizing the forms of Γ pdf's) gives us

$$.5\frac{\gamma^2\Gamma(2)^2}{(t_1+\gamma)^2(t_2+\gamma)^2}\,.$$

The odds ratio for equality over inequality is therefore

$$\frac{2(t_1+\gamma)^2(t_2+\gamma)^2}{\gamma(t_1+t_2+\gamma)^3} .$$

This odds ratio does not depend only on the statistic w we used in the previous part. It can be written as a function of w and $u = t_1 + t_2$, using the fact that $t_1 = u \cdot (1+w)/2$, $t_2 = u \cdot (1-w)/2$. When we do so, we see that because the numerator of the odds ratio is of order u^4, while the denominator is of order u^3, very large values of u tend to favor the equality hypothesis, given a fixed value for w. The intuition for this is that the prior we have chosen implies that the unconditional pdf for $u \mid w$, with α's integrated out, has much fatter tails under $\alpha_1 = \alpha_2$ than when we allow $\alpha_1 \neq \alpha_2$. This might reflect better performance for the odds ratio than for the test statistic of the previous section; but it might also reflect careless choice of priors.