STAT 215 – Midterm 1 NAME:

`Use extra sheets as required.  11pts per item.`

1. The following data come from a 1980 court case involving alleged applicant discrimination.

	rejected	not rejected	total
men	30	14	44
women	39	41	80

Denote with $x = 30$ and $y = 39$ the number of rejected men and women, respectively.

1a. Set up a model $y \sim p(y|\theta_y)$ and $x \sim p(x|\theta_x)$.

1b. Give the UMVUE for θ_x (justify your answer, i.e., demonstrate that the estimator which you use is in fact UMVUE).

1c. Set up a non-informative prior for θ_x and θ_y, assuming them *a priori* independent.

1d. Determine the posterior distribution $p(\theta_x, \theta_y|x, y)$.

1e. Give a normal approximation to the posterior $p(\theta_x, \theta_y|x, y)$.

1f. Using the normal approximation from part 1e, approximately evaluate the posterior probability $P(\theta_x > \theta_y|x, y)$ (If you do not have a normal table, you can write the answer in terms of normal c.d.f. values).

2. Suppose y_i, $i = 1, \ldots, n$, is a random sample from a $U(0, \theta)$ distribution, where θ is unknown, and suppose the prior distribution of θ is Pareto with parameters $\gamma > 0$ and $\lambda > 0$.

Note: The pareto p.d.f. is given by:

$$p(\theta|\lambda, \gamma) = \begin{cases} \frac{\lambda\gamma^\lambda}{\theta^{\lambda+1}} & \text{if } \theta \geq \gamma \\ 0, & \text{if } \theta < \gamma. \end{cases}$$

2a. Find the posterior distribution $p(\theta|y_1, \ldots, y_n)$.

2b. For an additional observation y_{n+1} we only know that $(y_{n+1} < z)$ for some z (a "censored" observation). Find the posterior $p(\theta|y_1, \ldots, y_n, z)$.

3. Assuming the regularity conditions hold, we know that $p(\theta|y)$ approaches normality as $n \to \infty$. In addition, if $\phi = f(\theta)$ is any one-to-one continuous transformation of θ, we can express the Bayesian inference in terms of ϕ and find that $p(\phi|y)$ approaches normality. But it is well known that a nonlinar transformation of a normal distirbution is no longer normal. How can both limiting normal distributions be valid? Outline an argument in less than 10 lines.