Table 1: *Number of respondents in each category for pre- and post-debate survey results*

Survey	Bush	Dukakis	No opinion/other	Total
pre-debate	294	307	38	639
post-debate	288	332	19	639

Homework Assignments #2

1. **Definition of k-parameter exponential family:** A probability density $p(y \mid \boldsymbol{\theta})$, is said to belong to the k-parameter exponential family if it is of the form

$$p(y \mid \boldsymbol{\theta}) = f(y) g(\boldsymbol{\theta}) \exp\left\{\sum_{j=1}^{k} c_j \phi_j(\boldsymbol{\theta}) h_j(y)\right\}$$

If $y = (y_1, \ldots, y_i, \ldots, y_n)$ is an independent and identically distributed sample of size n from the sampling distribution $p(y_i \mid \boldsymbol{\theta})$, then $\boldsymbol{t}_n(y_1, \ldots, y_n) = (n, \sum_{i=1}^{n} h_1(y_i), \ldots, \sum_{i=1}^{n} h_k(y_i))$ is a sequence of sufficient statistics.

Show that $p(y \mid \mu, \sigma^2) = N(y \mid \mu, \sigma^2)$ belongs to the k-parameter exponential family for $k = 2$;

Find $\phi_j(\boldsymbol{\theta}), h_j(y), c_j,\ j = 1, 2$, and $\boldsymbol{t}_n$;

Define the conjugate prior $p(\mu, \sigma^2)$;

Derive explicitly the posterior distribution $p(\mu, \sigma^2 \mid y)$.

2. Comparison of two multinomial experiments: on September 25, 1988, the evening of a Presidential campaign debate, ABC news conducted a survey of registered voters in the US; 639 persons were polled before the debate, and 639 different persons were polled after. The results are displayed in Table 1;

Assume that the surveys are independent samples from the population of registered voters. Model the data with two different multinomial distributions. For $j = 1, 2$, let α_j be the proportion of voters who preferred Bush, out of those who had a preference for either Bush or Dukakis at the time of survey j. Plot a histogram of the posterior density for $\alpha_2 - \alpha_1$. What is the posterior probability that there was a shift toward Bush?