Family Name: *Given Name:* *Student Number:*

UNIVERSITY OF TORONTO
Faculty of Arts and Science

DECEMBER EXAMINATIONS 2009

STA 437H1 F (plus STA 1005)

Duration - 3 hours

1	
2	
3	
4	
5	
6	
T	

No books, notes, or calculators are allowed.

The six questions are worth equal amounts.

Answer in the space provided; if you run out, use the back of a page (and point to where).

Except as noted, when the answer is a number, you must provide an actual number (eg, 1.5 or 3/2), not just a formula that could be evaluated to give this number.

Except as noted, you must explain how you obtained your answer to obtain full credit.

The following formulas may (or may not) be useful

Covariance of transformed random vector: $\text{Cov}(\mathbf{CX}) = \mathbf{C\Sigma_X C'}$

Probability density function for multivariate normal:

$$f(\mathbf{x}) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp(-(\mathbf{x}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)/2)$$

Conditional mean and covariance for multivariate normal:

$$\text{Mean of } \mathbf{X}_1 \text{ given } \mathbf{X}_2 = \mathbf{x}_2 = \mu_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \mu_2)$$

$$\text{Covariance of } \mathbf{X}_1 \text{ given } \mathbf{X}_2 = \mathbf{x}_2 = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$$

T^2 statistic for one sample: $T^2 = n(\overline{\mathbf{X}} - \mu_0)'\mathbf{S}^{-1}(\overline{\mathbf{X}} - \mu_0)$

The distribution of T^2 under the null hypothesis is $[(n-1)p/(n-p)]F_{p,n-p}$, which is approximately χ^2_p when $n-p$ and n/p are both large.

T^2 statistic for two samples, using pooled covariance estimate:

$$T^2 = ((\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) - \delta_0)'\,[(1/n_1 + 1/n_2)\,\mathbf{S}_{\text{pooled}}]^{-1}\,((\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) - \delta_0)$$

Here, $\mathbf{S}_{\text{pooled}} = ((n_1-1)\mathbf{S}_1 + (n_2-1)\mathbf{S}_2)\,/\,(n_1+n_2-2)$. The distribution of T^2 is $[(n_1+n_2-2)p/(n_1+n_2-p-1)]F_{p,n_1+n_2-p-1}$ under the null hypothesis that $\mu_1 - \mu_2 = \delta_0$. This distribution is approximately χ^2_p when n_1+n_2-p and $(n_1+n_2)\,/\,p$ are both large.

The factor analysis model: $\mathbf{X} = \mu + \mathbf{LF} + \epsilon$
$\mathbf{F} \sim N(0, \mathbf{I})$ and independently $\epsilon \sim N(0, \mathbf{\Psi})$, with $\mathbf{\Psi}$ diagonal.