iv. \( X_1, \ldots, X_k \sim N \text{Bin}(\theta, m) \)
- Number of trials it takes to get \( m \) successes, if each has success probability \( \theta \)
- Likelihood for one observation
  \[ L(\theta) = \left(\frac{X}{m-1}\right)^m \theta^m (1-\theta)^{X-m} \]
- Log likelihood for one observation
  \[ \ln(L) = \ln\left(\left(\frac{X}{m-1}\right)^m \theta^m (1-\theta)^{X-m}\right) + m \ln(\theta) + (X-m)\ln(1-\theta) \]
- Overall log likelihood
  \[ \sum_{j=1}^{n} \ln\left(\left(\frac{X_j}{m-1}\right)^m \theta^m (1-\theta)^{X_j-m}\right) + m \ln(\theta) + (\sum_{j=1}^{n} X_j - nm)\ln(1-\theta) \]
  
  \[ l' (\theta) = nm/\theta - (\sum_{j=1}^{n} X_j - nm)/(1-\theta) \]
  \[ l''(\theta) = \frac{nm}{\theta^2} - \frac{(\sum_{j=1}^{n} X_j - nm)}{(1-\theta)^2} \]

v. Bivariate Normal
- \( X_i \sim N(0,1), Y_i \mid X_i \sim N(\rho X_i, 1-\rho^2) \)
- \( l(\rho) = \)
  \[ \sum_{j=1}^{n} \left[ - \frac{X_j^2}{2} - \frac{(Y_j - \rho X_j)^2}{2(1-\rho^2)} - \ln(2\pi) - \frac{1}{2} \ln(1-\rho^2) \right] \]
- \( l'(\rho) = \)
- \( l''(\rho) = \)
- \( \rho^2 \) satisfies \( \rho(\rho^2 - 1) = 0 \)
- By invariance, \( \hat{\rho} = \tilde{X} \), unbiased.

\[ F: 10.8 \]

iv. Ensuring that our local max is a global max, possibly by checking
- that either \( L''(\theta) < 0 \ \forall \theta \)

i. Properties of Maximum Likelihood estimates?
- Are they unbiased? No, but almost...
- Are they consistent? No, but almost...
- Are they efficient? No, but almost...

j. For ind. observations,
  \[ f_{X_1, j, \ldots, X_n}(x_1, \ldots, x_n; \theta) = \prod_{j=1}^{n} f_{X_j}(x_j; \theta) \]
  
  and hence
  \[ L(\theta; X_1, \ldots, X_n) = \prod_{j=1}^{n} L(\theta; X_j) \]
  
  and
  \[ l(\theta; X_1, \ldots, X_n) = \sum_{j=1}^{n} l(\theta; X_j) \]

so the log likelihood for a collection of ind. random variables is the sum of the ind. likelihoods.

4. In general, m.o.m.e. = m.l.e. if density or mass function is of form \( \exp(c(\theta)x + b(\theta) + d(x)) \)