Mods Statistics, Sheet 1, HT 2012

1. Suppose $X_1, \ldots, X_n$ is a random sample from a distribution with mean μ and variance σ^2. Let $\overline{X} = \sum_{i=1}^n X_i/n$ and $S^2 = \sum_{i=1}^n (X_i - \overline{X})^2/(n-1)$ be the sample mean and variance.

(i) Find $E(\overline{X})$ and $\mathrm{var}(\overline{X})$.

(ii) Using $\sum(X_i - \overline{X})^2 = \sum\{(X_i - \mu) + (\mu - \overline{X})\}^2$ show that

$$\sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\overline{X} - \mu)^2.$$

By taking expectations show that $E(S^2) = \sigma^2$.

2. Let $X_1, \ldots, X_n$ be independent identically distributed random variables. Find the maximum likelihood estimators of the parameter θ for the following distributions. (In each case r is a known positive integer.)

(i) X_i has a binomial distribution with parameters r and θ.

(ii) X_i has a negative binomial distribution with probability mass function

$$f(x;\theta) = \binom{r+x-1}{x} \theta^r (1-\theta)^x, \quad x = 0, 1, 2, \ldots.$$

(iii) X_i has a gamma distribution with probability density function

$$f(x;\theta) = \frac{\theta^r}{(r-1)!} x^{r-1} e^{-\theta x}, \quad x > 0.$$

3. Suppose that in a population of twins, males (M) and females (F) are equally likely to occur and that the probability that twins are identical is θ. If twins are not identical, their sexes are independent.

(i) Show that $P(MM) = P(FF) = (1+\theta)/4$ and $P(MF) = (1-\theta)/2$.

(ii) Suppose that n twins are sampled. It is found that n_1 are MM, n_2 are FF, and n_3 are MF, but it is not known which twins are identical. Find the maximum likelihood estimator of θ.

4. Suppose X is a normal random variable with mean μ and variance σ^2.

(i) If a and b are constants, show that $aX + b$ has a normal distribution and find its mean and variance.

(ii) If $Z = (X - \mu)/\sigma$, deduce that $Z \sim N(0,1)$.

(iii) Using (ii) find $P(X < x)$ in terms of Φ, where Φ is the cumulative distribution function of a $N(0,1)$ random variable.

(iv) If $c > 0$ is a constant, show that $P(\mu - c\sigma < X < \mu + c\sigma)$ does not depend on μ or σ.

5. It is a fact, which you may assume, that if X and Y are independent and normally distributed random variables, then $X + Y$ is normally distributed.

Suppose $X_1, \ldots, X_n$ are independent normal random variables, X_i having mean μ_i and variance σ_i^2. If $a_1, \ldots, a_n$ are constants, show that $\sum_{i=1}^n a_i X_i$ is normally distributed and find its mean and variance.