Exercise 5

1. (i) Let $b_i = \alpha a_i + \beta$, $i = 1, \ldots, n$, and $\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$, $\bar{b} = \frac{1}{n}\sum_{i=1}^{n} b_i$. Show that $\bar{b} = \alpha\bar{a} + \beta$, and
$$\sum_{i=1}^{n}(b_i - \bar{b})^2 = \alpha^2 \sum_{i=1}^{n}(a_i - \bar{a})^2.$$
(Note that the sample variance of $\{b_i\}$ is independent of β.)
(ii) Let $\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$ and $\bar{b} = \frac{1}{n}\sum_{i=1}^{n} b_i$. Show that
$$\sum_{i=1}^{n}(a_i - c)(b_i - d) = \sum_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b}) + n(\bar{a} - c)(\bar{b} - d)$$
for any constants c and d, and
$$\sum_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b}) = \sum_{i=1}^{n} a_i b_i - n\bar{a}\bar{b}$$

2. Let $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ be two independent samples, where $X_1, \ldots, X_n$ are independent observations from $N(\mu_1, \sigma^2)$, and $Y_1, \ldots, Y_m$ are independent observations from $N(\mu_2, \sigma^2)$. Define
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,$$
$$\bar{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i, \quad s_y^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \bar{Y})^2.$$
Find the sampling distributions of (i) $\bar{X}-\bar{Y}$, (ii) $T \equiv (n-1)s_x^2+(m-1)s_y^2$, and (iii) $\frac{\bar{X}-\bar{Y}}{\sqrt{T}}\sqrt{\frac{nm(n+m-2)}{m+n}}$ when $\mu_1 = \mu_2$.

3. A finite population has N_a individuals of type A and N_b individuals of type B. The overall number N_a+N_b is equal to 100, but N_a is unknown. If a sample of 5 individuals is taken without replacement, write down the likelihood function for N_a when the observed sequence is (i) A, B, A, B, A, or (ii) A, A, B, B, A. Compare them with the likelihood function when the precise sequence of the outcomes is not available, but it is known that three elements are A and two others are B, and comment on your findings.

4. Given a random sample $Y_1, ..., Y_n$ from the following distributions, find a simple sufficient statistic using the factorisation criterion.
   (a) Poisson with mean λ
   (b) Normal with mean 0 and variance σ^2
   (c) Geometric
   (d) Uniform on (θ_1, θ_2) (Hint: $f_Y(y;\theta_1,\theta_2) = \frac{1}{\theta_2-\theta_1} I_{(\theta_1,\infty)}(y) I_{(-\infty,\theta_2)}(y)$.)
   (e) density $f_Y(y;\theta) = \exp(-(y-\theta))$ for $y > \theta$ (Use indicator functions, and take care with the factorisation)

