1. (5%) If the true value of the quantity being measured is denoted by x_0, the measurement, X, is modeled as

$$X = x_0 + \beta + \varepsilon$$

where β is the constant, or systematic, error and ε is the random component of the error; ε is a random variable with $E(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma^2$. An overall measure of the size of the measurement error that is often used is the mean squared error, (MSE), which is defined as $MSE = E[(X - x_0)^2]$. Find the MSE in terms of β and σ^2.

2. (13%)If $F(x)$ and $G(y)$ are one-dimensional cdfs, it can be shown that, for any α for which $|\alpha| \leq 1$,

$$H(x,y) = F(x)G(y)\{1 + \alpha[1 - F(x)][1 - G(y)]\}$$

is a bivariate cumulative distribution function.
Because $\lim_{x\to\infty} F(x) = \lim_{y\to\infty} G(y) = 1$, the marginal distributions are

$$H(x,\infty) = F(x), \quad H(\infty,y) = G(y).$$

As an example, we will construct bivariate distribution with marginals that are uniform on [0,1].

(i) (5%) If $\alpha = 1$, find $H(x,y)$.

(ii)(8%) If $\alpha = 1$, find $Cov(X,Y)$.

3. (26%)If gene frequencies are in equilibrium, the genotypes AA, Aa, and aa occur in population with frequencies $(1-\theta)^2$, $2\theta(1-\theta)$, and θ^2, according to the Hardy-Weinberg law. If we let X_1, X_2, and X_3 denote the counts in the three cells and let $n = X_1 + X_2 + X_3$,

   (i) (5%) Find the maximum likelihood estimator (MLE) of θ.

   (ii) (3%) Is the MLE of θ unbiased?

   (iii) (6%) Find an approximate 95% confidence interval for θ.

   (iv) (12%) In a sample from the Chinese population of Hong Kong in 1937, blood types occurred with the following frequencies, when M and N are erythrocyte antigens. Use the likelihood ratio test

	Blood Type			
	M	MN	N	Total
Frequency	342	500	187	1029

   to test whether this observed data is consistent with the Hardy-Weinberg law at $\alpha = 5\%$. ($\chi^2_{1;0.95} = 3.84$, $\chi^2_{2;0.95} = 5.99$).

4. (30%)Let X have the pdf

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

   Consider the random variable $Y = -2\log X$.

   (i) (4%) Find the cdf of Y.

   (ii) (6%) Find the moment generating function of Y.

(iii) (10%) Suppose the random sample $Y_1, Y_2, \ldots, Y_n$ from the distribution of Y. Let $Z_n = Y_1 + Y_2 + \cdots + Y_n$, find the limiting distribution of $W_n = \frac{Z_n - n}{\sqrt{2n}}$.

(iv) (10%) Find the limiting distribution of $\sqrt{n}(V_n - (\frac{1}{2})^{\frac{1}{4}})$, where $V_n = (\frac{Z_n}{\sqrt{2n}})^{\frac{1}{2}}$.

5. (18%) Let $X_1, X_2, \ldots, X_n$ be i.i.d with pdf

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}x} & 0 < x < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

(i) (4%) Find the minimum variance unbiased estimator, (MVUE), of θ.

(ii) (6%) When $n = 3$ and let $Z_1 = X_1 + X_2 + X_3$, $Z_2 = X_2 + X_3$, $Z_3 = X_3$, find the joint pdf of Z_1, Z_2, Z_3.

(iii) (8%) Find $E(\frac{Z_1}{3}|z_3)$ and $E[E(\frac{Z_1}{3}|z_3)]$.

6. (8%) Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution with mean μ and variance σ^2. To show that the sample variance converges in probability to σ^2, assume further that $E[X_1^4] < \infty$, so that $Var(S^2) < \infty$.