Statistics 700 Homework 0

Prerequisite

Due date: 6:00 pm (EST) Sept. 12, 2017

(1). **Coinflip Trivia.** Assume that you are randomly flipping a biased coin, the probability of getting a head 'H' is 0.4 and the probability of getting a tail 'T' is 0.6. The coin flips are independent of each other.

1. What is the probability of getting 'HHT' in the first three flips?
2. Conditioning on 'two of the first five coin flips are head', what is the probability that the first flip is 'T'? Conduct numerical simulations to verify your result.
3. If you earn \$ 10 for each 'H' and \$ 6 for each 'T', what is the expected reward and corresponding variance for $2, 3, 4, \ldots, n$ coin flips? Conduct numerical simulations to verify your result for 10 coinflips.

(2). **Exchangeability.** A generalization of iid random variables is *exchangeable* random variables, an idea due to deFinetti (1972). A discussion of exchangeability can also be found in Feller (1971). The random variables $X_1, \ldots, X_n$ are exchangeable if any permutation of any subset of them of size k ($k \leq n$) has the same distribution. In this exercise we will see an example of random variables that are exchangeable but not iid. Let $X_i|P \sim$ iid Bernoulli(P), $i = 1, \ldots, n$, and let $P \sim$ uniform$(0, 1)$.

1. Show that the marginal distribution of any k of the Xs is the same as

$$P(X_1 = x_1, \ldots, X_k = x_k) = \int_0^1 p^t(1-p)^{k-t}dp = \frac{t!(k-t)!}{(k+1)!},$$

   where $t = \sum_{i=1}^k x_i$. Hence, the Xs are exchangeable.

2. Show that, marginally,

$$P(X_1 = x_1, \ldots, X_n = x_n) \neq \prod_{i=1}^n P(X_i = x_i),$$

   so the distribution of the Xs is exchangeable but not iid.

(3). **Gamma-Poisson**. Given that $N = n$, the conditional distribution of Y is χ^2_{2n}. The unconditional distribution of N is Poisson(θ).

1. Calculate $E(Y)$ and $Var(Y)$ (unconditional moments).

2. Verify your results in 1 through simulations with $\theta = 1$: for $m = 10, 100, 1000$, simulate Y_i ($1 \leq i \leq m$), calculate the sample mean and sample variance. Try to visualize your results through repeated simulations.

3. Show that as $\theta \to \infty$, $(Y - E(Y))/\sqrt{Var(Y)} \to N(0, 1)$ in distribution.

4. Verify your results in 3 through simulations with $\theta = 1, 10, 100, \ldots$: for each θ, simulate Y_i ($1 \leq i \leq m$), make a histogram of $(Y - E(Y))/\sqrt{Var(Y)}$ and compare it with the density of a standard Gaussian distribution. Experiment on different values of m, what do you find as you increase m or θ?

(4). **Gaussian Example.** Assume that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are independently sampled from a bivariate normal distribution with parameters $(\mu_X, \mu_Y, \sigma^2_X, \sigma^2_Y, \rho)$.

1. Assume that $\rho \in (-1, 1)$ is known and $\sigma = \sigma_X = \sigma_Y$.

   (a) Derive the maximum likelihood estimators of (μ_X, μ_Y, σ^2).

   (b) Let $n = 100$ and $\sigma^2_X = \sigma^2_Y = 1, \rho = 0.3$. We can use two methods to form a 95% confidence set for $\theta = \mu_Y/\mu_X$: (1) the Fieller's theorem (Fieller 1954) and (2) the asymptotic distribution of $\widehat{\theta} = \sum_i Y_i / \sum_i X_i$ obtained from the Delta method. In your numerical simulations, set the true values $\mu_X = \mu_Y = 10$, and compare the 95% confidence sets from the two methods based on simulated data. What do you find from comparing the results through repeated simulations as you vary n?

2. Assume that $\rho \in (-1, 1)$ is unknown.

   (a) Consider a linear transformation of (X_i, Y_i)

   $$W_i = \frac{X_i - \mu_X}{\sigma_X} - \rho\frac{Y_i - \mu_Y}{\sigma_Y}, Z_i = \frac{Y_i - \mu_Y}{\sigma_Y}.$$

   What is the joint distribution of (W_i, Z_i)?

   (b) Derive the maximum likelihood estimator (MLEs) of ρ and show its consistency theoretically and numerically (through a series of repeated simulation studies, as $n \to \infty$, visualize your results).