

Statistical Inference 453/553. Spring 2003

Solutions HW1. Selected Problems

February 21, 2002

1. **Exercise 5.4.** (a) Given P the X 's are iid Bernoulli(p) random variables, $P[X_1 = x_1, \dots, X_k = x_k | p] = p^t(1-p)^{k-t}$ where $t = \sum_{i=1}^k x_i$. Since $P \sim U(0, 1)$ then,

$$P[X_1 = x_1, \dots, X_k = x_k] = \int_0^1 P[X_1 = x_1, \dots, X_k = x_k | p] dp = \int_0^1 p^t(1-p)^{k-t} dp.$$

The integrand in the last expression defines the kernel of a $Beta(t+1, k-t+1)$ pdf. Then

$$\int_0^1 p^t(1-p)^{k-t} dp = \frac{\Gamma(t+1)\Gamma(k-t+1)}{\Gamma(k+1)} = \frac{t!(k-t)!}{(k+1)!}.$$

(b) Applying part (a), we have that $P[X_i = x_i] = \frac{x_i!(1-x_i!)}{2!} = 1/2; x_i = 0, 1$. Therefore,

$$P[X_1 = x_1, \dots, X_k = x_k] = \frac{t!(k-t)!}{(k+1)!} \neq (1/2)^n$$

2. **Exercise 5.16** Define $Z_i = \frac{(X_i - i)}{i}; i = 1, 2, 3$ which are independent $N(0, 1)$ random variables.

(a) If $T = \sum_{i=1}^3 Z_i^2$, then T follows a chi-square distribution with 3 degrees of freedom (see Lemma 5.3.2).

(b) Now define $T = Z_1 / \sqrt{(Z_2^2 + Z_3^2)/2}$, by the calculations on page 223, T follows a t -distribution with 2 degrees of freedom.

(c) Now make $T = Z_1^2 / (Z_2^2 + Z_3^2)$. This is a ratio of independent random variables, a chi-square with one degree of freedom and a chi-square with 2 degrees of freedom. The resulting density is F with 1 and 2 degrees of freedom.

3. **Exercise 5.35** (a) This is a direct application of the CLT. Since the X 's are exponential(1) random variables, $E(\bar{X}) = 1$ and $Var(\bar{X}) = 1/n$. Then, $Z_n = \frac{(\bar{X}_n - 1)}{1/\sqrt{n}}$ converges in distribution to Z a $N(0, 1)$ random variable.

(b) Since Z is a $N(0, 1)$, differentiating $P(Z \leq x)$ gives the $N(0, 1)$ pdf, i.e. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. For the left side expression, notice that

$$P\left(\frac{\bar{X}_n - 1}{1/\sqrt{n}} \leq x\right) = P\left(\sum_{i=1}^n X_i \leq x\sqrt{n} + n\right)$$

The differentiation of this last expression is equal to $\sqrt{n}f_s(x\sqrt{n} + n)$ where $f_s(\cdot)$ denotes the pdf of $S = \sum_{i=1}^n X_i$. Since each X_i is an exponential(1) random variable, $S = \sum_{i=1}^n X_i$ has a Gamma(n,1) distribution. Then $f_s(y) = (1/\Gamma(n))y^{n-1}e^{-y}$, and by replacing y for $x\sqrt{n} + n$, we obtain the desired expression.

Finally, for $x = 0$, we have that

$$\frac{\sqrt{n} n^n}{(n-1)!n} e^{-n} \approx \frac{1}{\sqrt{2\pi}}$$

This expression implies Stirling's formula. Just realize that the denominator of the fraction of the right side is $n!$.

4. **Exercise 5.39** (a) Since $|h(x_n) - h(x)| < \epsilon$ whenever $|x_n - x| < \delta$ then,

$$Pr\{|X_n - X| < \delta\} \leq Pr\{|h(X_n) - h(X)| < \epsilon\}$$

Given that X_n converges in probability to X , then $\lim_{n \rightarrow \infty} P\{|X_n - X| < \delta\} = 1$. Combining these last 2 statements, we obtain

$$1 \leq \lim_{n \rightarrow \infty} Pr\{|h(X_n) - h(X)| < \epsilon\}$$

Since probabilities are always bounded above by one, $h(X_n)$ converges in probability to $h(X)$.

(b) The subsequence was introduced in class as $X_n(s) = s + I_{(0,1/n)}(s)$.

5. **Exercise 5.49** (a) Define $Z = -\log(U)$ with U following a $U(0,1)$. Then, $U = \exp(-Z)$. The density function of Z is

$$f_Z(z) = f_U(\exp(-z)) | -\exp(-z) | = \exp(-z)$$

so Z follows an exponential(1) distribution. Since U is uniform(0,1), $X = 1 - U$ is also a uniform(0,1). Then, by the previous argument, $Z = -\log(X)$ is an exponential(1) random variable.

(b) In this case, $U = e^X/(1 + e^X)$ and $dU/dX = e^X/(1 + e^X)^2$. Then,

$$f_X(x) = f_U(e^x/(1 + e^x)) e^x/(1 + e^x)^2 = e^x/(1 + e^x)^2$$

the logistic(0,1) distribution.

(c) Part (b) tells us that we may generate a logistic(0,1) random variable through the expression $X = \log(U/(1-U))$. Additionally, if X follows a logistic(0,1), $Z = \mu + \beta X$ follows a logistic(μ, β) distribution. We can generate Z with the expression $Z = \mu + \beta \log(U/(1-U))$.