Statistical Inference I **Prof. Chin-Tsang Chiang**

Lecture Notes 8

October 21, 2015 **Scribe: Wei-Chang Lee, Chi-Ning Chou**

Today we are going to talk about some examples of transformation in random variable and probability integral transformation.

0.1 Examples of transformation

0.1.1 Binomial distribution

Let X be a random variable with binomial distribution, then $f_X(x) = \begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x}\mathbf{1}_{0,1,\ldots,n}(x)$. Suppose $Y = n - X$, then Y is also a random variable.

0.1.2 Exponential

Let X be a random variable with exponential distribution, then $f_X(x) = \lambda e^{-\lambda x}\mathbf{1}_{(-\infty,\infty)}(x)$. Suppose $Y = X^\gamma$, $\gamma > 0$, then Y is also a random variable. Furthermore, the we can derive the distribution of Y as follow:

$$F_Y(y) = \int_{\{x:\ x^\gamma \le y\}} = \int_0^{y^{1/\gamma}} f_X(x)dx f_Y(y) = \frac{\lambda}{\gamma} y^{\frac{1}{\gamma}-1} e^{-\lambda y^{\frac{1}{\gamma}}} \mathbf{1}_{(0,\infty)}(y)$$

Remark 1 Power transformation v.s. natural logarithm transformation

Box and Cox introduced the power transformation:

$$\frac{X^\gamma - 1}{\gamma}$$

Moreover, we can see that as $\gamma \to 0$, the power transformation becomes natural logarithm transformation:

$$\ln X$$

These two transformations are similar but with different properties. In practice, we sometimes need to perform test to fit one of them to our model.

Remark 2 Degree of freedom

When using **linear model** plus **normal-family distribution**, degree of freedom refers to the rank of the power of our description variable. Most of the time, the degree of freedom is $n - k$, where n is the number of samples and k is the number of statistics we use.

Remark 3 Mean or median?

Mean and median are two different statistics to describe the central location of a data. We want to use them to estimate the population center. However, we might afraid of the existence of outliers so that the result will be biased. Thus, we can propose other methods to inference the central location as long as it's meaningful.

Remark 4 Box-Muller transformation: Generating normal distribution

Since the inverse of normal distribution cannot be written down in a close form, we cannot simply plugging uniformly distributed random variable to generate normal random variable. Instead, there's a method based on the intuition to calculate $\int_{-\infty}^{\infty} e^{-x^2/2} dx$ called Box-Muller transformation. Intuitively, we consider the polar coordinate (R, θ), and let

- $R^2 = X^2 + Y^2$, where X and Y are independent normal distribution.
- $\theta = 2\pi U_1$, where U_1 is a uniform distribution on [0,1].

As R^2 is actually a chi-square distribution with degree of freedom 2, we can write it as $R^2 = -2 \ln U_2$. As a result, we can generate a normal distributed random variable as

$$Z = R \cos\theta = \sqrt{-2 \ln U_2} \cos(2\pi U_1)$$

Remark 5 Poisson approximation v.s. normal approximation

The close form of binomial distribution involves lots of binomial terms. As a result, when n is large, it's computationally inefficient to directly compute the cumulative distribution or pmf from the close form. We need some computationally efficient approximation. There are two kinds of approximations for binomial random variable: Poisson and normal. What's the difference? Here, we state two approximations and compare their intuitions.

Poisson approximation

$$\begin{aligned}
\binom{n}{k} p^k (1-p)^{n-k} &= \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} p^k (1-p)^{n-k} \\
(\text{let } \lambda = np) &= \frac{1}{k!} \frac{n \cdot (n-1) \cdots (n-k+1)}{n \cdot n \cdots n} \lambda^k (1 - \frac{\lambda}{n})^{n-k} \\
(n \text{ large}) &\approx \frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}$$

As p is small and n is large, Poisson approximation can give a good results.

Normal approximation As n grows large, by central limit theorem, we have

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{(k/n-p)^2}{2}}$$