3 Special Distributions

ENSY SILVER[1]

Tuesday 4th August, 2020

[1]Thanks to my family, my friend and freedom.

1 Poisson distribution

1.1 Poisson limit

Suppose X is a binomial random variable, where

$$P(X = k) = p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

If $n \to \infty$ and $p \to 0$ in such a way that $\lambda = np$ remains constant, then

$$\lim_{\substack{n\to\infty \\ p\to 0 \\ np=\text{constant}}} P(X = k) = \frac{e^{-np}(np)^k}{k!}$$

The proof is trivial, we omit it.

1.2 Poisson distribution

Theorem 1.1. *The random variable X is said to have a Poisson distribution if*

$$p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

where λ is a positive constant. Also, for any Poisson random variable, $E(X) = \lambda$ and $\operatorname{Var}(X) = \lambda$.

Remark. *The book provided an intuitive sight to Poisson model. Consider a time interval of length T, and an event which has expected occurence $E(X)$, then we divide the interval into nonoverlapping subintervals with equivalent length. As the length of a subinterval goes to zero, we have the pdf of the occurence.*

From the *Poisson distribution*, we can also deduce the distribution of length of intervals between two consecutive occurence.

Theorem 1.2. *Suppose a series of events satisfying the Poisson model are occurring at the rate of λ per unit time. Let the random variable Y denote the interval between consecutive events. Then Y has the exponential distribution.*

$$f_Y(y) = \lambda e^{-\lambda y}, \quad y > 0$$

2 Normal distribution

2.1 Nature of normal distribution

Determining probabilities of certain outcomes of a random variable becomes more complicated when we are given only the mean or both the mean and the variance. However, while exact probabilitiesare impossible to find, bounds on the probability can be derived. One such bound is given by Markov ' s inequality

Lemma 2.1 (**Markov's inequality**). *If X is a random variable that takes only nonnegative values, then for any value $a > 0$,*

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Proof. The proof is remarkable, hence I write it here. Define a new random variable I such that

$$I = \begin{cases} 1, & \text{if } X \geq a \\ 0, & \text{otherwise} \end{cases}$$

We have the condition that $I < \frac{X}{a}$. Then, it is natural that $E[I] < E[\frac{X}{a}]$. Reviewing the definition of I, we find that $E[I] = P(X \geq a)$. So $P(X \geq a) = E[I] \leq \frac{E[X]}{a}$. Hence the lemma is proved. □

Directly applying lemma 2.1, we have another lemma.

Lemma 2.2 (**Chebyshev's Inequality**). *If X is a finite random variable with finite mean μ and variance σ^2, then for any value $k > 0$,*

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

The lemma 2.2 also leads to the **weak law of large numbers**.

Theorem 2.3 (**weak law of large numbers**). *Let $X_1, X2, \cdots, X_n$ be a sequence of independent and identically distributed random variables, each having finite mean $E[X_i] = \mu$ and variance σ^2. Then, for any $\epsilon > 0$,*

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) \longrightarrow 0 \quad \text{as} \quad n \to 0$$

Before the core part of this section, we introduce a lemma.

Lemma 2.4. *Let $Z_1, Z_2, \cdots$ be a sequence of random variables having distribution functions F_{Z_n} and mgfs M_{Z_n} where $n \geq 1$. Furthermore, let Z be a random variable having distribution function F_Z and mgf M_Z. If $M_{Z_n}(t) \to M_Z(t)$ for all t, then $F_{Z_n}(t) \to F_Z(t)$ for all t at which $F_Z(t)$ is continuous.*

Now we reach the milestone of the section, **central limit theorem**.

Theorem 2.5. *Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables with mean μ and variance σ^2. Then the distribution of $\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}}$ tends to the standard normal as $n \to \infty$.*

$$P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2} 2 \, dx \quad \text{as} \quad n \to \infty$$

The proof is at [1]. Now we finally get the physically nature of normal distribution, that is, the summation of large amount of independent and indentical random variables tends to be normally distributed.

2.2 Normal distribution and its properties

Here, we give the formal definition of normal distribution.

Theorem 2.6. *A random variable Y is said to be normally distributed with mean μ and variance σ^2 if*

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) \quad -\infty < y < \infty$$

The symbol $Y \sim \mathcal{N}(\mu, \sigma^2)$ will sometimes be used to denote the fact that Y has a normal distribution with mean μ and variance σ^2.

Remark. *We use $\exp(x)$ to replace e^x inorder to make the index larger in paper.*

Now, we introduce the first property, the relationship between mgf and normal distribution.

Theorem 2.7. *For a normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$, the mgf is*

$$M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

The combination of two normal distributed random variables is also critical.

Theorem 2.8. *Let Y_1 be a normally distributed random variable with mean μ_1 and variance σ_1^2, and let Y_2 be a normally distributed random variable with mean μ_2 and variance σ_2^2. Define $Y = Y_1 + Y_2$. If Y_1 and Y_2 are independent, Y is normally distributed with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.*

The proof is directly derived from the combination of mgf.

3 Geometric distribution

Consider a series of independent trials, each having one of two possible outcomes, success or failure. Let $p = P(\text{Trial ends in success})$. Define the random variable X to be the trial at which the *first success occurs*. The pdf of X is

$$p_X(k) = (1-p)^{k-1} p$$

The probability model above is called a **geometric distribution**. The basic properties of geometric distribution are mean, variance and mgf.

Theorem 3.1. *Let X have a geometric distribution with $p_X(k) = (1-p)^{k-1}p$, where $k = 1, 2, \cdots$. Then*

1. $M_X(t) = \frac{pe^t}{1-(1-p)e^t}$.
2. $E(X) = \frac{1}{p}$.
3. $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

Applying mgf, we can easily deduce the theorem 3.1.

4 Negative binomial distribution

In this section, we expand the notion of geometric distribution. Imagine waiting for the r^{th} (instead of the first) success in a series of independent trials, where each trial has a probability of p of ending in success. Let the random variable X denote the trial at which the r^{th} success occurs. Then

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad k = r, r+1, \cdots$$

Now, we introduce the formal definition of **negative binomial distributed**.

Theorem 4.1. *Let X have a negative binomial distribution with $p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$, where $k = r, r+1, \cdots$. Then*

1. $M_X(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]^r$.
2. $E(X) = \frac{r}{p}$.
3. $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

5 Gamma distribution

Theorem 5.1. *Suppose that Poisson events are occurring at the constant rate of λ per unit time. Let the random variable Y denote the waiting time for the r^{th} event. Then Y has pdf $f_Y(y)$, where*

$$f_Y(y) = \frac{\lambda^r}{(r-1)!} y^{r-1} e^{-\lambda y} \tag{1}$$

Now, we consider to expand the 1 for real number r, hence we need to introduce **gamma function** Γ, which is the extension of $r!$. More details about gamma function will be introduced in Complex Analysis notes.

Theorem 5.2. *For any real number $r > 0$, the gamma function of r is denoted be $\Gamma(r)$, where*

$$\Gamma(r) = \int_0^\infty y^{r-1} e^{-y} \, dy$$

We shall notice that Γ has intriguing properties.

Theorem 5.3. *Let $\Gamma(r) = \int_0^\infty y^{r-1} e^{-y} \, dy$ for any real number $r > 0$. Then*

1. $\Gamma(1) = 1$.
2. $\Gamma(r) = (r-1)\Gamma(r-1)$.
3. $\Gamma(r) = (r-1)!$ *when* $r \in \mathbb{Z}$.

So, we replace the $(r-1)!$ in 1 by $\Gamma(r)$, and get the formal definition of **gamma distribution**.

Theorem 5.4. *Given real numbers $r > 0$ and $\lambda > 0$, the random variable Y is said to have the gamma pdf with parameters r and Γ if*

$$f_Y(y) = \frac{\lambda^r}{\Gamma(r)} y^{r-1} e^{-\lambda y}$$

Then, we turn to some basic properties of gamma distribution.

Theorem 5.5. *Suppose that Y has a gamma pdf with parameters r and λ. Then*

1. $E(Y) = r/\lambda$.

2. $\mathrm{Var}(Y) = r/\lambda^2$.

The combination of gamma distribution is computable.

Theorem 5.6. *Suppose U has the gamma pdf with parameters r and λ, V has the gamma pdf with parameters s and λ, and U and V are independent. Then $U + V$ has a gamma pdf with parameters $r + s$ and λ.*

The mgf of gamma distribution is simpler the its pdf.

Theorem 5.7. *If Y has a gamma pdf with parameters r and λ, then*

$$M_Y(t) = \left(1 - \frac{t}{\lambda}\right)^{-r}$$

6 Review

In this part, the difference between continuous random variables and discrete random variables plays an important role. When counting occurences for a discrete random variable, we use summation. But for continuous random variables, we calculate the pdf by differentiating the probability function.

References

[1] Vlad Krokhmal. *Introductory Probability and the Central Limit Theorem.* `http://www.math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Krokhmal.pdf`. July 2011.