北京交通大學

Special Distributions

Chapter 5 of Probability and Statistics

Yiping Cheng
ypcheng@bjtu.edu.cn

10-Jul-2017

School of Electronic and Information Engineering
Beijing Jiaotong University

Chapter Contents

Special Discrete Distributions

Bernoulli Distributions

Definition: A random variable X is said to have the *Bernoulli distribution* with parameter p $(0 \leq p \leq 1)$ if

$$P(X = 1) = p, \quad P(X = 0) = 1 - p.$$

$$E(X) = 1 \times p + 0 \times (1 - p) = p,$$
$$E(X^2) = 1^2 \times p + 0^2 \times (1 - p) = p,$$
$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1 - p).$$

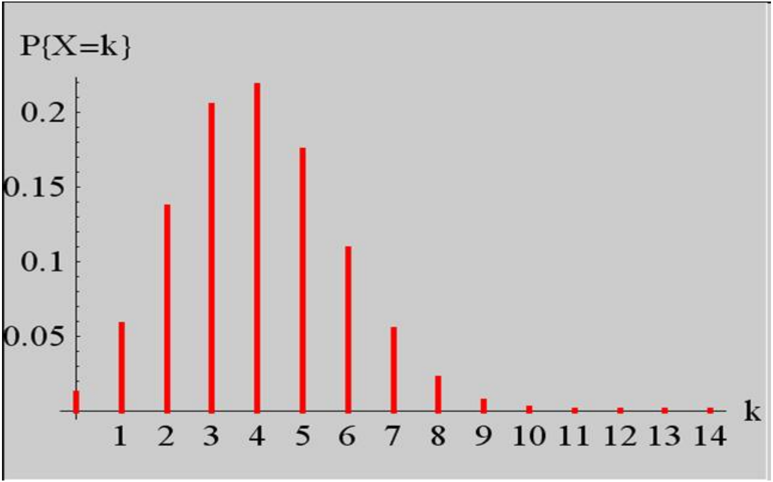
Binomial Distributions

Definition [Binomial Distribution]: A random variable X is said to have the binomial distribution with parameters n and p (n positive integer, $0 \leq p \leq 1$) if

$$P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{for } x = 0, 1, \ldots, n \\ 0 & \text{otherwise.} \end{cases}$$

Remark

The Bernoulli distributions are special cases of the binomial distributions ($n = 1$).


P{X=k}
0.2
0.15
0.1
0.05
1 2 3 4 5 6 7 8 9 10 11 12 13 14
k

Mean and Variance of Binomial Distributions

Theorem

Let $X \sim \text{binomial}(n, p)$. Then

$$E(X) = np \text{ and } \text{Var}(X) = np(1-p).$$

Proof.

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-1-l)!} p^l (1-p)^{n-1-l} = np[p + (1-p)]^{n-1} = np.$$

$$E[X(X-1)] = \sum_{k=1}^{n} k(k-1) \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!} p^k (1-p)^{n-k} = n(n-1)p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-k)!} p^{k-2} (1-p)^{n-k}$$

$$= n(n-1)p^2 \sum_{l=0}^{n-2} \frac{(n-2)!}{l!(n-2-l)!} p^l (1-p)^{n-2-l} = n(n-1)p^2 [p + (1-p)]^{n-2} = n(n-1)p^2.$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p).$$

□

Additive Property of Binomial Distributions

Theorem: If $X_1, \ldots, X_k$ are independent random variables and if X_i has the binomial distribution with parameters n_i and p $(i = 1, \ldots, k)$, then the sum $X_1 + \cdots + X_k$ has the binomial distribution with parameters $n_1 + \cdots + n_k$ and p.

Remark

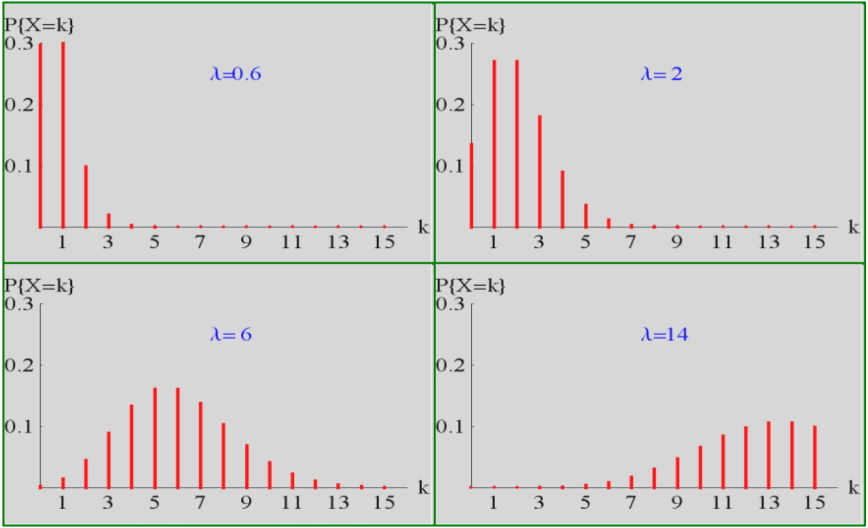
- This theorem is a generalization of a similar result in Chapter 3. Its proof is omitted because that will involve the concept of characteristic functions which is beyond the scope of this course.
- It is essential that the component RVs be independent and their distributions have the same parameter p.

Poisson Distributions

The Poisson distributions are introduced as limiting distributions of the binomial distributions when n is very large but np is moderate. For example, suppose there is a shop in a town of 10000 people, and in a particular hour, each person in the town, independently, has 0.005 probability to come to the shop. The number of people that come to the shop in that hour is therefore a random variable having the binomial distribution with parameters $n = 10000$ and $p = 0.005$. We will later show that this distribution can be approximated by the Poisson distribution with parameter $10000 \times 0.005 = 50$.

Definition [Poisson Distribution]: A random variable X is said to have the Poisson distribution with parameter λ $(\lambda \geq 0)$ if

$$P(X = x) = \begin{cases} \frac{\lambda^x}{x!}e^{-\lambda} & x = 0, 1, \ldots \\ 0 & \text{otherwise.} \end{cases}$$


P{X=k}
0.3
0.2
0.1
λ=0.6
λ= 2
λ= 6
λ=14
k
1 3 5 7 9 11 13 15

Mean and Variance of Poisson Distributions

Theorem: Let $X \sim \text{Poisson}(\lambda)$. Then

$$E(X) = \lambda \text{ and } \text{Var}(X) = \lambda.$$

Proof.

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda e^{\lambda} e^{-\lambda} = \lambda.$$

$$E[X(X-1)] = \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} e^{-\lambda} = \lambda^2 e^{\lambda} e^{-\lambda} = \lambda^2.$$

$$\begin{aligned} \text{Var}(X) &= E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2 \\ &= \lambda^2 + \lambda - \lambda^2 = \lambda. \quad \square \end{aligned}$$

$\square$

Additive Property of Poisson Distributions

Theorem: If $X_1, \ldots, X_k$ are independent random variables and if X_i has the Poisson distribution with parameters λ_i $(i = 1, \ldots, k)$, then the sum $X_1 + \cdots + X_k$ has the Poisson distribution with parameter $\lambda_1 + \cdots + \lambda_k$.

Remark

This theorem can be easily proved using characteristic functions or moment generating functions. However, these two concepts are beyond this course.

Poisson Approximation to Binomial

Theorem: Suppose $np_n = \lambda$ for $n = 1, 2, \ldots$. Then

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Proof.

Denote $f_n(k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}$. Then

$$\begin{aligned} f_n(k) &= \frac{n(n-1)\cdots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!} (\frac{\lambda}{n})^k (1-\frac{\lambda}{n})^{n-k} \\ &= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k(1-\frac{\lambda}{n})^k} (1-\frac{\lambda}{n})^n \\ &= \frac{\lambda^k}{k!} (1-\frac{\lambda}{n})^n (1-\frac{\lambda}{n-\lambda})(1-\frac{\lambda-1}{n-\lambda})\cdots(1-\frac{\lambda-k+1}{n-\lambda}). \end{aligned}$$

Since $\lim_{n\to\infty}(1-\frac{\lambda}{n-\lambda})(1-\frac{\lambda-1}{n-\lambda})\cdots(1-\frac{\lambda-k+1}{n-\lambda}) = 1$, we have

$$\lim_{n\to\infty} f_n(k) = \lim_{n\to\infty} \frac{\lambda^k}{k!}(1-\frac{\lambda}{n})^n = \frac{\lambda^k}{k!} e^{-\lambda}.$$

□

Poisson Approximation to Binomial – Example

Diseased People. Suppose that in a large population the proportion of people who have a certain disease is 0.01. Determine the probability that in a group of 200 people at least four people will have the disease.

Solution Method 1: Let X be the number of people in the group who have the disease. Then $X \sim \text{binomial}(n = 200, p = 0.01)$, and

$$P(X \geq 4) = 1 - \sum_{i=0}^{3} P(X = i) = 1 - \binom{200}{0} 0.01^0 \cdot 0.99^{200} - \binom{200}{1} 0.01^1 \cdot 0.99^{199}$$

$$- \binom{200}{2} 0.01^2 \cdot 0.99^{198} - \binom{200}{3} 0.01^3 \cdot 0.99^{197} = 0.141965965555...$$

Poisson Approximation to Binomial – Example

Diseased People. Suppose that in a large population the proportion of people who have a certain disease is 0.01. Determine the probability that in a group of 200 people at least four people will have the disease.

Solution Method 1: Let X be the number of people in the group who have the disease. Then $X \sim \text{binomial}(n = 200, p = 0.01)$, and

$$P(X \geq 4) = 1 - \sum_{i=0}^{3} P(X = i) = 1 - \binom{200}{0} 0.01^0 \cdot 0.99^{200} - \binom{200}{1} 0.01^1 \cdot 0.99^{199}$$

$$- \binom{200}{2} 0.01^2 \cdot 0.99^{198} - \binom{200}{3} 0.01^3 \cdot 0.99^{197} = 0.141965965555...$$

Solution Method 2: Approximately let us assume $X \sim \text{Poisson}(\lambda = np = 2)$. Hence

$$P(X \geq 4) = 1 - \sum_{i=0}^{3} P(X = i) = 1 - (\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!})e^{-2} = 0.1428765...$$

The error is quite small.

Introducing Negative Binomial Distributions

Sampling until a Fixed Number of Successes. Suppose that an infinite sequence of independent Bernoulli trials with probability of success p are available. Determine the p.f. of the random number X of failures that occur before the rth success.

Solution: The event $X = k$ occurs if and only if exactly $r - 1$ successes occur among the first $r - 1 + k$ trials and the $(r + k)$th trial is successful. Since all trials are independent, it follows that

$$\begin{aligned} P(X = k) &= \binom{r-1+k}{k} p^{r-1}(1-p)^k \cdot p \\ &= \binom{r-1+k}{k} p^r (1-p)^k. \end{aligned}$$

Negative Binomial Distributions

Definition [Negative Binomial Distribution]: A random variable X is said to have the negative binomial distribution with parameters r and p (r positive integer, $0 < p \leq 1$) if

$$P(X = x) = \begin{cases} \binom{r-1+x}{x} p^r (1-p)^x & \text{for } x = 0, 1, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Example

Defective Parts. A machine produces parts that can be either good or defective. Assume that the parts are good or defective independently of each other with p for all parts. An inspector observes the parts produced by this machine until she sees four defectives. Let X be the number of good parts observed by the time that the fourth defective is observed. Then $X \sim \text{negative\_binomial}(4, p)$.

Geometric Distributions

The geometric distributions are special cases of the negative binomial distributions ($r = 1$). Usually a geometric distribution will be used to model the number of failures until the first success.

Definition [Geometric Distribution]: A random variable X is said to have the geometric distribution with parameter p $(0 < p \leq 1)$ if it has the negative binomial distribution with parameters $r = 1$ and p, which means

$$P(X = x) = \begin{cases} p(1-p)^x & \text{for } x = 0, 1, \dots \\ 0 & \text{otherwise.} \end{cases}$$

Mean and Variance of Negative Binomial Distributions

The following theorem verifies that the previously given negative binomial p.f. is indeed a p.f. and give the expectation and variance of the negative binomial distributions. Its proof is not given because they will require the *generalized binomial theorem* which is beyond the scope of this course.

Theorem: Suppose that r is a positive integer, $0 < p \leq 1$, then

1. $\sum_{k=0}^{\infty} \binom{r-1+k}{k} p^r (1-p)^k = 1.$
2. Let $X \sim \text{negative\_binomial}(r, p)$. Then

$$E(X) = \frac{r(1-p)}{p} \text{ and } \mathrm{Var}(X) = \frac{r(1-p)}{p^2}.$$

Additive Property of Negative Binomial Distributions

Theorem: If $X_1, \ldots, X_k$ are independent random variables and if X_i has the negative binomial distribution with parameters r_i and p $(i = 1, \ldots, k)$, then the sum $X_1 + \cdots + X_k$ has the negative binomial distribution with parameters $r_1 + \cdots + r_k$ and p.

Remark

- This theorem is a generalization of Theorem 5.5.2 (P298). Its proof is omitted because that will involve the concept of characteristic functions which is beyond the scope of this course.
- It is essential that the component RVs be independent and their distributions have the same parameter p.

Poisson Approximation to Negative Binomial

Theorem: When $r \to \infty$ and $r(1-p_r) = \lambda$, the negative binomial distribution with parameters r and p_r tends to the Poisson distribution with parameter λ, i.e.

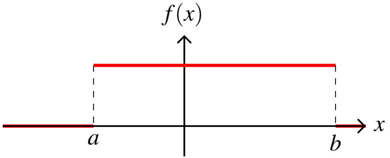
$$\lim_{r\to\infty} \binom{r-1+k}{k} p_r^r (1-p_r)^k = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Special Continuous Distributions

Uniform Distributions

Definition [Uniform Distribution]: A random variable X is said to have the uniform distribution on interval $[a, b]$ $(a < b)$ if X has a continuous distribution with the following p.d.f.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$



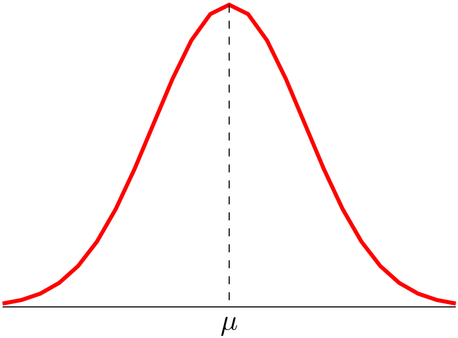
$$E(X) = \frac{a+b}{2},$$

$$\text{Var}(X) = \int_a^b (x - \frac{a+b}{2})^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}.$$

Normal Distributions

Definition [Normal Distribution]: A random variable X is said to have the normal distribution with mean μ and variance σ^2 ($\sigma > 0$) if X has a continuous distribution with the following p.d.f.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$



Verification of Normal PDF

Lemma: Let $a > 0$, then $\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$.

Proof.

Let $I = \int_{-\infty}^{\infty} e^{-ax^2} dx$, then $I = \int_{-\infty}^{\infty} e^{-ay^2} dy$, and

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dxdy = \int_0^{2\pi} d\theta \int_0^{\infty} e^{-ar^2} rdr \\ &= \pi \int_0^{\infty} e^{-ar^2} dr^2 = \pi(-\frac{e^{-\infty}}{a} - \frac{-e^0}{a}) = \frac{\pi}{a}. \end{aligned}$$

Hence $I = \sqrt{\frac{\pi}{a}}$. $\square$

Theorem [Verification of Normal PDF]:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1.$$

Mean and Variance of Normal Distributions

Theorem: Let $X \sim \text{normal}(\mu, \sigma^2)$. Then

$$E(X) = \mu \text{ and } \mathrm{Var}(X) = \sigma^2.$$

Proof.

$$E(X) = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{z+\mu}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}} dz$$

$$= \underbrace{\int_{-\infty}^{\infty} \frac{z}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}} dz}_{=0, \text{ and absolutely integrable}} + \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}} dz = \mu.$$

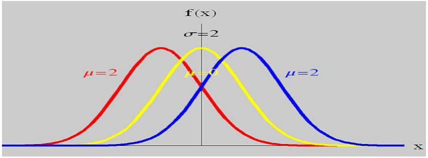
$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{\sigma^2 t^2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

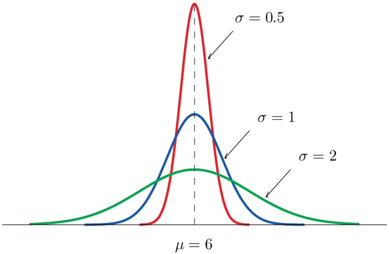
$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt\right) = \frac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = \sigma^2.$$

□

Curve Shape of Normal PDF

- The curve is symmetric with respect to $x = \mu$. It is bell-shaped. Changing μ only shifts the curve.
- The shape of the curve is mainly determined by σ. Bigger σ means fatter curve.





Invariance of Normality under Linear Transformations

Theorem: If X has the normal distribution with mean μ and variance σ^2 and if $Y = aX + b$, where a and b are given constants and $a \neq 0$, then Y has the normal distribution with mean $a\mu + b$ and variance $a^2\sigma^2$.

Proof.

Let f_Y be the p.d.f. of Y. Since $Y = r(X) = aX + b$, we have $r^{-1}(y) = \frac{y-b}{a}$.

Thus, $f_Y(y) = f_X[r^{-1}(y)]|\frac{dr^{-1}(y)}{dy}| = f_X[\frac{y-b}{a}]|\frac{1}{|a|}$

$$= \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{(y-a\mu-b)^2}{2a^2\sigma^2}},$$

which means $Y \sim \text{normal}(a\mu + b, a^2\sigma^2)$. □

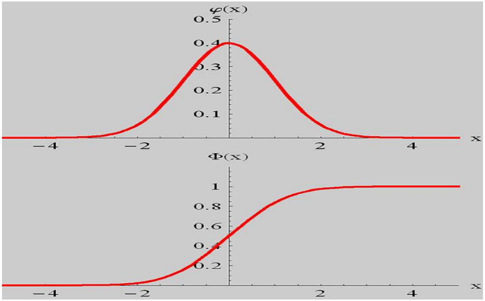
Corollary: If $X \sim \text{normal}(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim \text{normal}(0, 1)$.

The Standard Normal Distribution

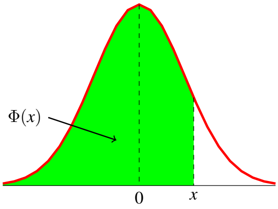
Definition: The normal distribution with mean 0 and variance 1 is called the *standard normal distribution*.

The p.d.f. of the standard normal distribution is usually denoted by the symbol φ, and the c.d.f. is denoted by the symbol Φ. Thus,

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi(x) = \int_{-\infty}^{x} \varphi(t)dt.$$



A Consequence of Even PDF



Theorem: $\Phi(-x) = 1 - \Phi(x)$.

Proof.

$$\Phi(-x) = \int_{-\infty}^{-x} \varphi(t)dt = \int_{u=\infty}^{u=x} \varphi(-u)d(-u) = \int_{x}^{\infty} \varphi(-u)du.$$

Since $\varphi(u) = \varphi(-u)$,

$$\Phi(-x) = \int_{x}^{\infty} \varphi(u)du = 1 - \Phi(x).$$

$\square$

Finding Probabilities for Normal Distributions

Let X have the normal distribution with mean μ and variance σ^2. Let F be the c.d.f. of X. Then $Z = (X - \mu)/\sigma$ has the standard normal distribution, and for all x,

$$F(x) = P(X \leq x) = P(Z \leq \frac{x-\mu}{\sigma}) = \Phi(\frac{x-\mu}{\sigma}).$$

Example

Suppose that X has the normal distribution with mean 5 and standard deviation 2. Determine the value of $P(1 < X < 8)$.

Solution: Let $Z = (X - 5)/2$, then Z has the standard normal distribution and

$$P(1 < X < 8) = P(\frac{1-5}{2} < \frac{X-5}{2} < \frac{8-5}{2}) = P(-2 < Z < 1.5)$$
$$= P(Z < 1.5) - P(Z \leq -2) = \Phi(1.5) - \Phi(-2) = \Phi(1.5) - [1 - \Phi(2)].$$

From the Φ-table, it is found that $\Phi(1.5) = 0.9332$ and $\Phi(2) = 0.9773$. Therefore, $P(1 < X < 8) = 0.9105$.

Finding Quantiles for Normal Distributions

Theorem: Let $X \sim \text{normal}(\mu, \sigma^2)$ and F be the c.d.f. of X. Then

$$\begin{aligned} F^{-1}(p) &= \mu + \sigma\Phi^{-1}(p), \\ \Phi^{-1}(p) &= -\Phi^{-1}(1-p). \end{aligned}$$

Proof

1. Follows directly from

$$F[\mu + \sigma\Phi^{-1}(p)] = P[X \leq \mu + \sigma\Phi^{-1}(p)]$$

$$= P[\frac{X-\mu}{\sigma} \leq \Phi^{-1}(p)] = \Phi[\Phi^{-1}(p)] = p.$$

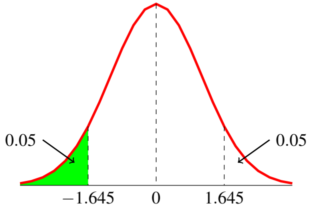
2. Follows directly from

$$\Phi[-\Phi^{-1}(1-p)] = 1 - \Phi[\Phi^{-1}(1-p)] = 1 - (1-p) = p.$$

Finding Quantiles for Normal Distributions – Example

Suppose that X has the normal distribution with mean 1.329 and standard deviation 0.4844. Find the 0.05 quantile of X.

Solution: We have $\Phi^{-1}(p) = -\Phi^{-1}(1-p)$.



So to find $\Phi^{-1}(0.05)$, look up 0.95 in the $\Phi(x)$ column of the Φ-table to find $x = 1.645$ and conclude that $\Phi^{-1}(0.05) = -1.645$. The 0.05 quantile of X is then
$F^{-1}(0.05) = 1.329 + 0.4844 \times (-1.645) = 0.5322$.

Additive Property of Normal Distributions

Theorem: If random variables $X_1, \ldots, X_k$ are independent, X_i has the normal distribution with mean μ_i and variance σ_i^2 ($i = 1, \ldots, k$), and $a_1, \ldots, a_k$ and b are constants for which at least one of $a_1, \ldots, a_k$ is not 0, then $a_1X_1 + \cdots + a_kX_k + b$ has the normal distribution with mean $a_1\mu_1 + \cdots + a_k\mu_k + b$ and variance $a_1^2\sigma_1^2 + \cdots + a_k^2\sigma_k^2$.

This theorem can be proved using convolution of p.d.fs or characteristic functions. However, the proof is a bit complicated so it is omitted here.

Example – Determining a Sample Size

Definition [Sample and Sample Mean]: Suppose that $X_1, \ldots, X_n$ are independent random variables all having the same distribution. Then they are collectively called a random sample from that distribution. n is the sample size. The average of these n random variables, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$, is called the sample mean.

Let $X_1, \ldots, X_n$ be a random sample from the distribution normal$(\mu, 9)$. What is the minimum sample size for which $P(|\bar{X} - \mu| \leq 1) \geq 0.95$?

Example – Determining a Sample Size

Definition [Sample and Sample Mean]: Suppose that $X_1, \dots, X_n$ are independent random variables all having the same distribution. Then they are collectively called a random sample from that distribution. n is the sample size. The average of these n random variables, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$, is called the sample mean.

Let $X_1, \dots, X_n$ be a random sample from the distribution normal$(\mu, 9)$. What is the minimum sample size for which $P(|\bar{X} - \mu| \le 1) \ge 0.95$?

Solution: It follows from the additive property that $\bar{X} \sim$ normal$(\mu, \frac{9}{n})$. Let $Z = \frac{\bar{X}-\mu}{3/\sqrt{n}}$. Then

$$P(|\bar{X} - \mu| \le 1) = P(|Z| \le \frac{\sqrt{n}}{3}) = \Phi(\frac{\sqrt{n}}{3}) - \Phi(\frac{-\sqrt{n}}{3}) = 2\Phi(\frac{\sqrt{n}}{3}) - 1.$$

Since it is required that $P(|\bar{X} - \mu| \le 1) \ge 0.95$, we have $2\Phi(\frac{\sqrt{n}}{3}) - 1 \ge 0.95$, i.e. $\Phi(\frac{\sqrt{n}}{3}) \ge 0.975$, and $\frac{\sqrt{n}}{3} \ge \Phi^{-1}(0.975) = 1.96$, and then $n \ge 34.6$. The sample size must be at least 35.

Multivariate Normal Distributions

Definition [Multivariate Normal Distribution]: The random variables $X_1, \ldots, X_n$ are said to have the n-variate normal distribution with mean $\mu = (\mu_1, \ldots, \mu_n)^T$ and covariance matrix Σ (Σ is positive definite) if $X_1, \ldots, X_n$ have a continuous joint distribution with the following p.d.f.

$$f(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} e^{-\frac{(\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)}{2}}$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T$.

Bivariate Normal Distributions

Bivariate normal distributions are special cases of multivariate normal distributions when $n = 2$. Let

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < \rho < 1$, which ensures Σ positive definite. Then the joint p.d.f. of X_1 and X_2 is

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2} \exp\{\frac{-1}{2(1-\rho^2)}$$

$$[(\frac{x_1-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2}) + (\frac{x_2-\mu_2}{\sigma_2})^2]\}.$$

Properties of Bivariate Normal Distributions

Suppose that X_1, X_2 have the bivariate normal distribution with mean $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ and covariance matrix $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$. Then

1. $X_i \sim \text{normal}(\mu_i, \sigma_i^2)$ for $i = 1, 2$.
2. $\text{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2$ and $\rho(X_1, X_2) = \rho$.
3. X_1 and X_2 are independent if and only if $\rho = 0$.
4. **[Linear Combination of Bivariate Normals]** Let $Y = a_1X_1 + a_2X_2 + b$ where a_1, a_2, b are constants. Then

$$Y \sim \text{normal}(a_1\mu_1 + a_2\mu_2 + b, \quad a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\rho\sigma_1\sigma_2).$$

Fundamental Theorem of Multivariate Normal Distributions

Suppose that $X_1, \ldots, X_n$ have the multivariate normal distribution with mean μ and covariance matrix Σ. If random vector $\mathbf{Y} = (Y_1, \ldots, Y_k)^T$ is obtained by

$$\mathbf{Y} = \mathbf{AX} + \mathbf{b},$$

where $\mathbf{A}$ is an $k \times n$ *full row-rank* constant matrix, $\mathbf{X} = (X_1, \ldots, X_n)^T$, and $\mathbf{b} = (b_1, \ldots, b_k)^T$ is a constant column vector, then

$Y_1, \ldots, Y_k$ have the multivariate normal distribution with

mean $\mathbf{A}\mu + \mathbf{b}$ and covariance matrix $\mathbf{A}\Sigma\mathbf{A}^T$.

Further Properties of Multivariate Normal Distributions

Suppose that $X_1, \ldots, X_n$ have the multivariate normal distribution with mean μ and covariance matrix Σ. Then

1. $X_i \sim \text{normal}(\mu_i, \Sigma_{ii})$ for $i = 1, \ldots, n$.
2. $\text{Cov}(X_i, X_j) = \Sigma_{ij}$.
3. Any subset of $X_1, \ldots, X_n$ also have a multivariate normal distribution.
4. Any subset of $X_1, \ldots, X_n$ are mutually independent if and only if they are pairwise uncorrelated.

The Gamma Function

Definition [Gamma Function]: For each positive number α, let the value $\Gamma(\alpha)$ be defined by the following integral:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx.$$

Theorem: $\Gamma(\alpha)$ exists as a finite number for any $\alpha > 0$.

Proof.

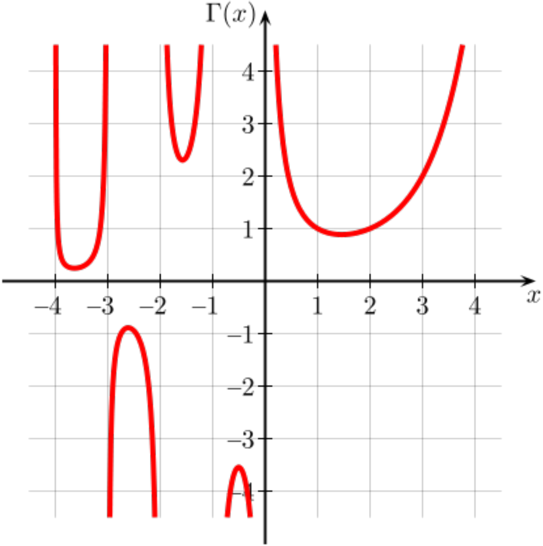
It is easy to see that there is some $h \geq e$ such that for all $x \geq h$, $\frac{x}{\ln(x)} \geq 2(\alpha - 1)$, and thus $x^{\alpha-1}e^{-x} \leq e^{\frac{-x}{2}}$. We then have

$$\Gamma(\alpha) = \int_0^1 x^{\alpha-1}e^{-x}dx + \int_1^h x^{\alpha-1}e^{-x}dx + \int_h^\infty x^{\alpha-1}e^{-x}dx$$

$$\leq \underbrace{\int_0^1 x^{\alpha-1}dx}_{=1/\alpha} + \underbrace{\int_1^h x^{\alpha-1}e^{-x}dx}_{\text{finite}} + \underbrace{\int_h^\infty e^{\frac{-x}{2}}dx}_{=2e^{\frac{-h}{2}}}.$$

$\square$

Curve of the Gamma Function



Properties of the Gamma Function

Theorem: For all $\alpha > 0$, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.

Proof.

$$\Gamma(\alpha+1) = \int_0^\infty x^\alpha e^{-x} dx = \int_{x=0}^{x=\infty} -x^\alpha de^{-x}$$

$$= -x^\alpha e^{-x}|_{x=0}^{x=\infty} + \int_{x=0}^{x=\infty} e^{-x} dx^\alpha = \int_0^\infty \alpha e^{-x} x^{\alpha-1} dx$$

$$= \alpha\Gamma(\alpha).$$

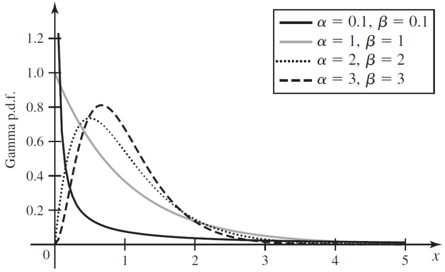
□

Corollary: For all positive integer n, $\Gamma(n) = (n-1)!$

Gamma Distributions

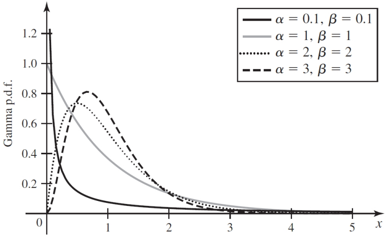
Definition [Gamma Distribution]: Let α and β be positive numbers. A random variable X is said to have the Gamma distribution with parameters α and β if X has a continuous distribution with the following p.d.f.

$$f(x) = \begin{cases} \frac{\beta}{\Gamma(\alpha)}(\beta x)^{\alpha-1}e^{-\beta x} & \text{for } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$



Curve Shape of Gamma PDF

- For $\alpha < 1$, the p.d.f. is strictly decreasing from $f(0) = \infty$ to $f(\infty) = 0$.
- For $\alpha = 1$, the p.d.f. is strictly decreasing from $f(0) = \beta$ to $f(\infty) = 0$.
- For $\alpha > 1$, the p.d.f. has a peak at $x = \frac{\alpha-1}{\beta^2}$. Before the peak, it is strictly increasing and after the peak, it is strictly decreasing.



Mean and Variance of Gamma Distributions

Theorem: Let $X \sim \text{Gamma}(\alpha, \beta)$. Then

$$E(X) = \frac{\alpha}{\beta} \text{ and } \text{Var}(X) = \frac{\alpha}{\beta^2}.$$

Proof.

$$E(X) = \int_0^\infty x \frac{\beta}{\Gamma(\alpha)}(\beta x)^{\alpha-1} e^{-\beta x} dx = \int_0^\infty \frac{1}{\Gamma(\alpha)}(\beta x)^{\alpha} e^{-\beta x} dx$$

$$= \int_0^\infty \frac{1}{\beta\Gamma(\alpha)}(\beta x)^{\alpha} e^{-\beta x} d\beta x = \frac{1}{\beta\Gamma(\alpha)}\Gamma(\alpha+1) = \frac{1}{\beta\Gamma(\alpha)}\alpha\Gamma(\alpha) = \frac{\alpha}{\beta}.$$

$$E(X^2) = \int_0^\infty x^2 \frac{\beta}{\Gamma(\alpha)}(\beta x)^{\alpha-1} e^{-\beta x} dx = \int_0^\infty \frac{1}{\beta\Gamma(\alpha)}(\beta x)^{\alpha+1} e^{-\beta x} dx$$

$$= \int_0^\infty \frac{1}{\beta^2\Gamma(\alpha)}(\beta x)^{\alpha+1} e^{-\beta x} d\beta x = \frac{1}{\beta^2\Gamma(\alpha)}\Gamma(\alpha+2)$$

$$= \frac{1}{\beta^2\Gamma(\alpha)}(\alpha+1)\alpha\Gamma(\alpha) = \frac{(\alpha+1)\alpha}{\beta^2}.$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{\alpha}{\beta^2}.$$

$\square$

Additive Property of Gamma Distributions

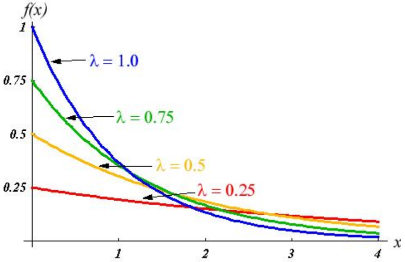
Theorem: If $X_1, \ldots, X_k$ are independent random variables and if X_i has the Gamma distribution with parameters α_i and β ($i = 1, \ldots, k$), then the sum $X_1 + \cdots + X_k$ has the Gamma distribution with parameters $\alpha_1 + \cdots + \alpha_k$ and β.

Exponential Distributions

The exponential distributions are special cases of the Gamma distributions where the parameter $\alpha = 1$.

Definition [Exponential Distribution]: Let $\beta > 0$. A random variable X is said to have the exponential distribution with parameter β if X has a continuous distribution with the following p.d.f.

$$f(x) = \begin{cases} \beta e^{-\beta x} & \text{for } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$



Memoryless Property of Exponential Distributions

Theorem: Let X have the exponential distribution with parameter β, and let $t > 0$. Then for every $h > 0$,

$$P(X \geq t + h | X \geq t) = P(X \geq h).$$

Proof.

For each $t > 0$,

$$P(X \geq t) = \int_t^{\infty} \beta e^{-\beta x} dx = e^{-\beta t}.$$

Hence, for each $t > 0$ and each $h > 0$,

$$P(X \geq t + h | X \geq t) = \frac{P(X \geq t + h)}{P(X \geq t)} = \frac{e^{-\beta(t+h)}}{e^{-\beta t}} = e^{-\beta h} = P(X \geq h).$$

$\square$

Homework for Chapter 5

- P280: 4, 8
- P296: 2, 4, 12, 13
- P301: 2, 6
- P315: 2, 5, 10, 11
- P325: 4, 6, 8
- P345: 2, 3, 4, 6, 13