STA2001 Probability and Statistics (I)

Lecture 11

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Review

Mathematical Expectations $E[g(X)] = \int_{\overline{S}} g(x)f(x)dx$

1. $[g(X) = X]$: Mean of X, $E[X] = \int_{\overline{S}} xf(x)dx$
2. $[g(X) = (X - E[X])^2]$: Variance of X,

$$Var[X] = E[(X - E[X])^2] = \int_{\overline{S}} (x - E[X])^2 f(x)dx$$

3. $[g(X) = X^r]$, Moments of X:

$$E[X^r] = \int_{\overline{S}} x^r f(x)dx$$

4. $[g(X) = e^{tX}]$: mgf, if there exists $h > 0$, such that

$$M(t) = E[e^{tX}] = \int_{\overline{S}} e^{tx} f(x)dx, \quad -h < t < h \text{ for some } h > 0$$

$$M^{(r)}(0) = E[X^r], E[X] = M'(0), \quad Var[X] = M''(0) - (M'(0))^2$$

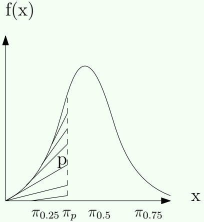
Review

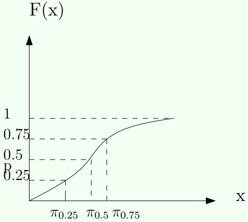
Definition[(100p)th percentile]

It is a number π_p such that the area under $f(x)$ to the left of π_p is p. That is

$$p = \int_{-\infty}^{\pi_p} f(x)dx = F(\pi_p)$$

The 50th percentile is called the median. The 25th and 75th percentiles are called the first and third quantiles, respectively. The median is also called the 2nd quantile

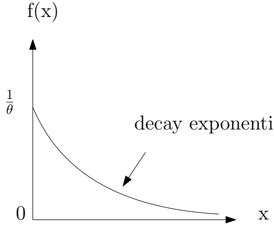




Review

- Exponential distribution with parameter $\theta = \frac{1}{\lambda}$:
  X, the waiting time until the first occurrence in an approximate Poisson process with parameter $\lambda > 0$ and its pdf takes the form of

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x \geq 0, \theta > 0$$



Mean and Variance:

$$E[X] = \theta, Var[X] = \theta^2$$

Mgf:

$$M(t) = \frac{1}{1 - t\theta}, \quad t < \frac{1}{\theta}$$

Poisson Distribution

Let X describe the number of occurrences of some events in a unit interval with

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \cdots, \quad E[X] = Var[X] = \lambda$$

For an interval with length T, which should be treated as a new "unit interval", the number of occurrences Y has $E[Y] = \lambda T$ and thus its pmf is

$$f(y) = \frac{(\lambda T)^y e^{-\lambda T}}{y!}, \quad y = 0, 1, \cdots$$

Poisson Distribution

Let X describe the number of occurrences of some events in a unit interval with

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \cdots, \quad E[X] = Var[X] = \lambda$$

For an interval with length T, which should be treated as a new "unit interval", the number of occurrences Y has $E[Y] = \lambda T$ and thus its pmf is

$$f(y) = \frac{(\lambda T)^y e^{-\lambda T}}{y!}, \quad y = 0, 1, \cdots$$

Then for $\alpha = 1, 2, \ldots,$

$$\begin{aligned} P(Y < \alpha) &= \sum_{k=0}^{\alpha-1} \frac{(\lambda T)^k e^{-\lambda T}}{k!} \\ &= P(\{\text{the number of occurrence smaller than } \alpha \\ &\quad \text{in the interval with length T}\}) \end{aligned}$$

Gamma distribution

1. Description: Consider an APP. We are interested in the waiting time until the αth occurrence, $\alpha = 1, 2, \ldots$.
2. Define the waiting time by W. Then our goal is to derive the pdf of W.

$$\text{Idea: } \begin{cases} 1.\text{derive cdf of W}, F(w) \\ 2. f(w) = F'(w) \end{cases}$$

$F(w) = P(W \leq w)$

Assume that the waiting time is nonnegative. Then,

$$F(w) = 0, \quad \text{for } w < 0.$$

For $w \geq 0$,

$$F(w) = P(W \leq w) = 1 - P(W > w)$$

$P(W > w) = P(\{\text{number of occurrences in } [0, w] \text{ smaller than } \alpha\})$

Gamma distribution

= Probability that in [0, w], there will be at most α-1 occurences

$$P(W > w) = P(\{\text{number of occurrences in } [0, w] \text{ smaller than } \alpha\})$$

$$= \sum_{k=0}^{\alpha-1} \frac{(\lambda \overset{(T)}{\overset{\uparrow}{w}})^k e^{-\lambda w}}{k!}$$

$$f(w) = F'(w) = \frac{\lambda^\alpha w^{\alpha-1}}{(\alpha-1)!} e^{-\lambda w}, w > 0.$$

pdf of this form is said to be of the Gamma type and W is said to have Gamma distribution.

The waiting time until the αth occurrence in the APP, has a Gamma distribution with parameters α and λ.

Gamma Function

With the so-called Gamma function, we can obtain a more general definition of Gamma distribution with $\alpha > 0$.

Definition[Gamma function (generalized factorial)]

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \quad t > 0$$

$$\Gamma(t) = -y^{t-1}e^{-y}\Big|_0^\infty + \int_0^\infty (t-1) y^{t-2} e^{-y} dy = (t-1)\Gamma(t-1)$$

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) \\ &= \cdots = (n-1)\cdots 2\Gamma(1) = (n-1)!(\Gamma(1) = 1)\end{aligned}$$

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \quad t > 0 = \int_0^\infty -y^{t-1} \, de^{-y}$$

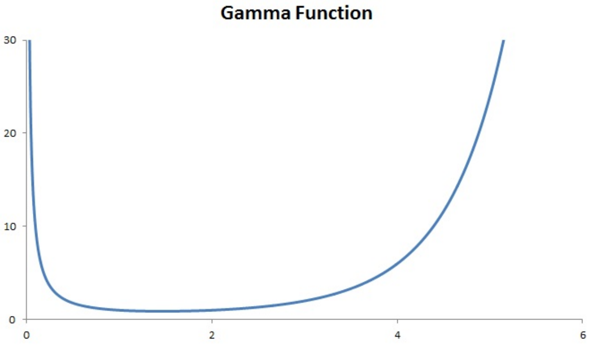
(t>1) $= -y^{t-1} e^{-y} \Big|_0^\infty + \int_0^\infty (t-1) y^{t-2} e^{-y} dy$

$$\Longrightarrow \Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \quad t > 0 = (t-1) \int_0^\infty y^{t-2} e^{-y} dy$$

$$\Longrightarrow \Gamma(t) = (t-1)\Gamma(t-1) \ , \quad t \geq 1$$

Gamma Function

With the so-called Gamma function, we can obtain a more general definition of Gamma distribution with $\alpha > 0$.



Gamma Distribution

Definition

A RV X has a Gamma distribution if its pdf is

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\theta}}, \quad x \geq 0, \alpha > 0, \theta > 0,$$

where θ and α are the two parameters.

- Gamma pdf $f(x)$ is well-defined pdf (by the definition of $\Gamma(\alpha)$)
- mgf (exercise 3.2-7)

$$M(t) = \frac{1}{(1-\theta t)^{\alpha}}, \quad t < \frac{1}{\theta}$$

$$E[x] = \alpha\theta, \quad Var[X] = \alpha\theta^2$$

A special case: when $\alpha = 1$, Gamma distribution reduces to exponential distribution.

Chi-square Distribution

Definition

Let X have a Gamma distribution with $\theta = 2$, $\alpha = \frac{r}{2}$, r is an integer. The pdf of X is

$$f(x) = \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}, \quad x > 0$$

Then X has chi-square distribution with degrees of freedom r, and denoted by $X \sim \chi^2(r)$

$$E[X] = \alpha\theta = \frac{r}{2} \cdot 2 = r \quad Var[X] = \alpha\theta^2 = \frac{r}{2} \cdot 2^2 = 2r$$

$$\text{Mgf}: M(t) = (1-2t)^{-\frac{r}{2}}, \quad t < \frac{1}{2}$$

Note: The interpretation of Chi-square distribution is deferred.

Remark

Chi-square distribution plays an important role in statistics, the tables of cdf of chi-square distribution are given

$$F(x) = P(X \leq x) = \int_0^x f(t)dt.$$

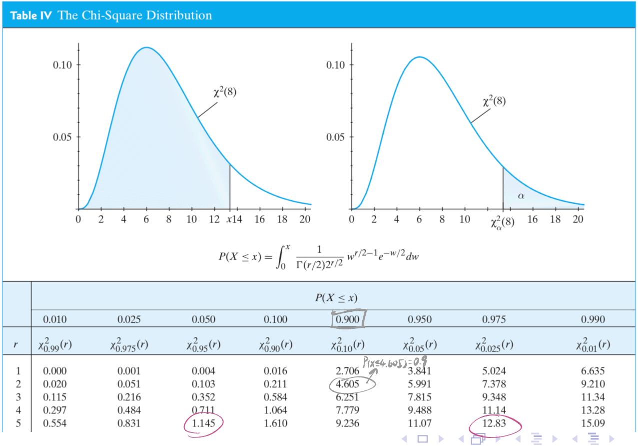


Table IV The Chi-Square Distribution

$$P(X \leq x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2} dw$$

	$P(X \leq x)$							
	0.010	0.025	0.050	0.100	$\boxed{0.900}$	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09

$P(X \leq 4.605) = 0.9$

Example 2

Let X have a chi-square distribution with $r = 5$ degrees of freedom. Then using table IV in Appendix B on page 501, we have

$$P(1.145 \leq X \leq 12.83) = F(12.83) - F(1.145)$$

Example 2

Let X have a chi-square distribution with $r = 5$ degrees of freedom. Then using table IV in Appendix B on page 501, we have

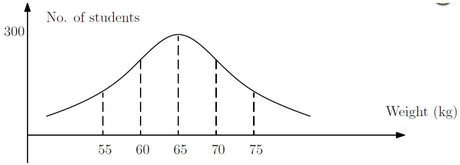
$$\begin{aligned} P(1.145 \leq X \leq 12.83) &= F(12.83) - F(1.145) \\ &= 0.975 - 0.05 = 0.925. \end{aligned}$$

3.3 Normal Distribution

Description

When observed over a large population, many things of interests have a “bell-shaped” relative frequency distribution.

- Weight of male students in CUHKsz
- Height
- TOFEL,IELTS test score



Normal Distribution

Definition

A continuous RV X is said to be normal or Gaussian if it has a pdf of the form

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}), \quad -\infty < x < \infty$$

where μ and σ^2 are two parameters characterizing the normal distribution. Briefly, $X \sim N(\mu, \sigma^2)$

pdf of Normal Distribution

$f(x)$ is a well-defined pdf

1. $f(x) > 0$ for all x.
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

We will prove $\int_{-\infty}^{\infty} f(x)dx = 1$ shortly, if time permits.

mgf, mean, variance

Assume $X \sim N(\mu, \sigma^2)$

$$M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx$$

$$e^{tx} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} = \exp\left\{-\frac{1}{2\sigma^2}[x^2 - 2(\mu + \sigma^2 t)x + \mu^2]\right\}$$

mgf, mean, variance

Consider

$$x^2 - 2(\mu + \sigma^2 t)x + \mu^2 = [x - (\mu + \sigma^2 t)]^2 - 2\mu\sigma^2 t - \sigma^4 t^2$$

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left[x - (\mu + \sigma^2 t)\right]^2\right) dx \\ \cdot \exp\left(\frac{-2\mu\sigma^2 t - \sigma^4 t^2}{-2\sigma^2}\right)$$

mgf, mean, variance

Recall that

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = 1, \text{independent of } \mu$$

Therefore,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}[x-(\mu+\sigma t)]^2\right) dx = 1$$

$$M(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

mgf, mean, variance

$$M(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) \implies M(0) = 1;$$

$$M'(t) = (\mu + \sigma^2 t)\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) \implies M'(0) = \mu$$

$$M''(t) = \sigma^2\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) + (\mu + \sigma^2 t)^2\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

$$\implies M''(0) = \mu^2 + \sigma^2$$

mgf, mean, variance

Recall that

$$E[X] = M'(0) = \mu$$

$$Var[X] = E[X^2] - (E[X])^2$$

$$= M''(0) - M'(0)^2 = \sigma^2$$

For $X \sim N(\mu, \sigma^2)$,

$$E[X] = \mu, \quad Var[X] = \sigma^2$$

The two parameters μ, σ^2 are the mean and variance, respectively.

Proof of $\int_{-\infty}^{\infty} f(s)ds = 1$

Let

$$I = \int_{-\infty}^{\infty} f(s)ds = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(s-\mu)^2}{\sigma^2}\right) ds \overset{?}{=} 1$$

Take a coordinate change $x = \frac{s-\mu}{\sigma}$

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \overset{?}{=} 1.$$

Since $I > 0$, then if $I^2 = 1$, then $I = 1$.

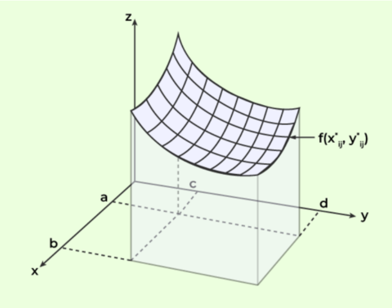
$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy$$

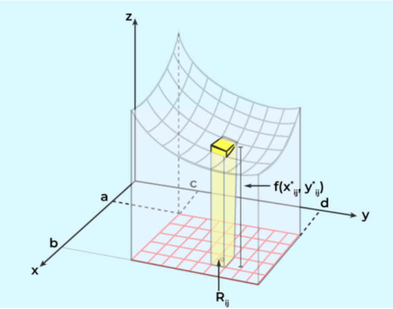
Geometric interpretation of double integral

For $f(x, y) \geq 0$

$$\text{Volume} = \iint_A f(x, y)dxdy$$

The double integral calculates the volume under the surface $f(x, y)$ over the definition domain of $f(x, y)$ or any region of interest A.





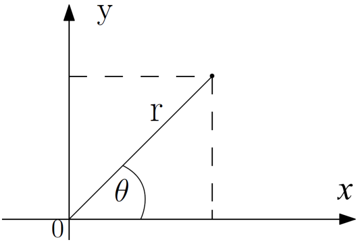
Geometric interpretation of double integral in polar coordinate system

Take a coordinate change

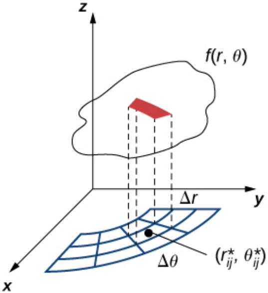
$$x = r\cos\theta$$

$$y = r\sin\theta$$

(polar coordinate)



The double integral calculates the volume under the surface $f(x, y)$ over the definition domain of $f(x, y)$ or any region of interest.



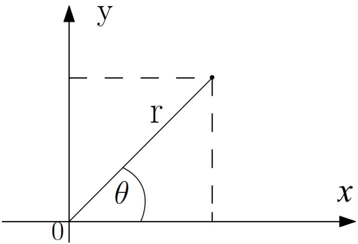
Geometric interpretation of double integral in polar coordinate system

Take a coordinate change

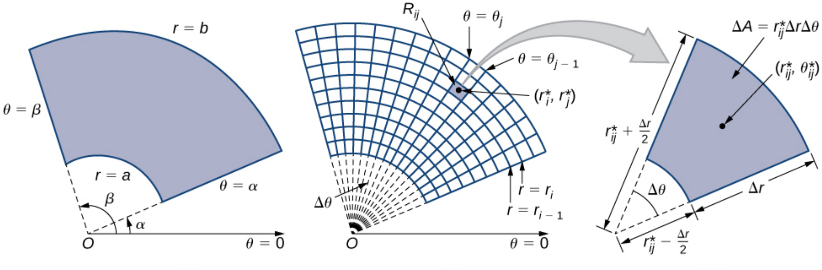
$$x = r\cos\theta$$

$$y = r\sin\theta$$

(polar coordinate)



The problem lies in to how to calculate the area of small trapezoid.



(a) (b) (c)

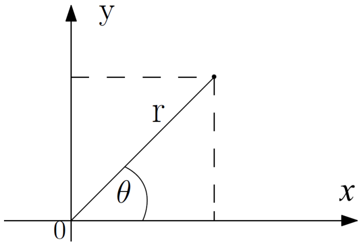
Proof of $\int_{-\infty}^{\infty} f(s)ds = 1$

Take a coordinate change

$$x = r\cos\theta$$

$$y = r\sin\theta$$

(polar coordinate)



$$I^2 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} r dr d\theta = \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_0^{\infty} e^{-\frac{r^2}{2}} d\frac{r^2}{2}$$
$$= \frac{1}{2\pi}\cdot 2\pi \cdot (-1)\cdot e^{-\frac{r^2}{2}}\big|_0^{\infty} = 1$$