STA2001 Probability and Statistics (I)

Lecture 12

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Review

Definition

X, the waiting time until the αth occurrence, and its pdf takes the form of

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

where $\theta > 0$ and $\alpha > 0$ are the two parameters,

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

- $\alpha = 1$, exponential distribution.
- $\theta = 2$, $\alpha = \frac{r}{2}$, r is an integer, chi square distribution (r is called the degrees of freedom).

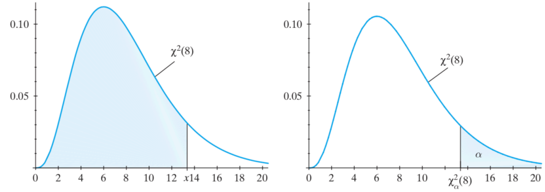
Review

Be able to calculate probabilities of events by looking up tables.

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	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

Review

Definition

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

A RV X is said to be normal or Gaussian if its pdf is

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

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

- $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) dx = 1$
- $M(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right), \quad t \in R$
- $E(X) = \mu, \quad Var(X) = \sigma^2$

Example 1, page 115

A RV X has its pdf in the form of

$$f(x) = \frac{1}{\sqrt{32\pi}} \exp\left[-\frac{(x+7)^2}{32}\right], -\infty < x < \infty$$

Example 1, page 115

A RV X has its pdf in the form of

$$f(x) = \frac{1}{\sqrt{32\pi}} \exp\left[-\frac{(x+7)^2}{32}\right], -\infty < x < \infty$$

$$\Leftrightarrow \quad X \sim N(-7, 16)$$

$$\Leftrightarrow E(X) = -7, Var(X) = 16$$

$$\Leftrightarrow M(t) = \exp(-7t + 8t^2).$$

Standard Normal Distribution

Y is said to be a standard normal distribution if

$$Y \sim N(0,1) \Leftrightarrow \text{ its pdf } f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

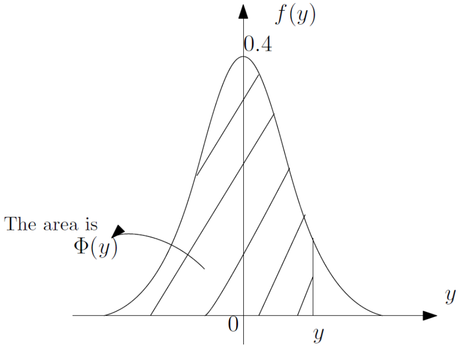
Its cdf

$$\Phi(y) = P(Y \le y) = \int_{-\infty}^{y} f(z)dz = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

Due to the symmetry of $f(y)$, $\Phi(-y) = 1 - \Phi(y)$, for any y

pdf of $N(0,1)$

$$Y \sim N(0,1) \Leftrightarrow \text{ its pdf } f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

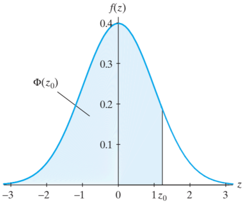


Due to the symmetry of $f(y)$, $\Phi(-y) = 1 - \Phi(y)$, for any y

pdf of $N(0,1)$

Values of $\Phi(y)$ for values of $y \geq 0$ are in Appendix B (page 502).

Table Va The Standard Normal Distribution Function



$$P(Z \leq z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

$$\Phi(-z) = 1 - \Phi(z)$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177

pdf of $N(0, 1)$

Values of $\Phi(y)$ for values of $y \geq 0$ are in Appendix B (page 502).

1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Example 2, page 116

$Z \sim N(0, 1)$ Then compute

$$P(Z \leq 1.24) = \Phi(1.24) = 0.8925$$

$$P(1.24 \leq Z \leq 2.37) = \Phi(2.37) - \Phi(1.24)$$

Example 2, page 116

$Z \sim N(0,1)$ Then compute

$$P(Z \leq 1.24) = \Phi(1.24) = 0.8925$$

$$P(1.24 \leq Z \leq 2.37) = \Phi(2.37) - \Phi(1.24)$$

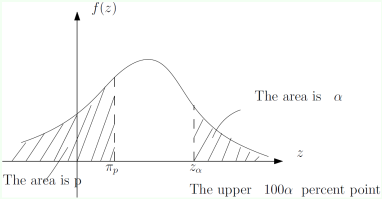
$$= 0.9911 - 0.8925 = 0.0986$$

$$P(-2.37 \leq Z \leq -1.24) = P(1.24 \leq Z \leq 2.37) = 0.0986$$

The upper 100α percent point

Definition

The number z_α such that $P(Z \geq z_\alpha) = \alpha$.



$P(X \leq \pi_p) = p$, π_p is 100pth percentile.

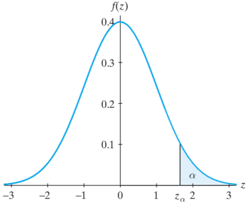
Note:

$$\begin{aligned} P(Z < z_\alpha) &= 1 - P(Z \geq z_\alpha) \\ &= 1 - \alpha \end{aligned}$$

So z_α is the $100(1-\alpha)$th percentile

The upper 100α percent point

Table Vb The Standard Normal Right-Tail Probabilities



$P(Z > z_\alpha) = \alpha$

$P(Z > z) = 1 - \Phi(z) = \Phi(-z)$

z_α	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823

Example 3, page 117

$Z \sim N(0,1)$, Find $Z_{0.0125}$. That is

$$P(Z \geq z_{0.0125}) = 0.0125$$

Example 3, page 117

$Z \sim N(0, 1)$, Find $Z_{0.0125}$. That is

$$P(Z \geq z_{0.0125}) = 0.0125$$

check the table $\Rightarrow z_{0.0125} = 2.24$

What about $z_{0.05}$ and $z_{0.025}$?

Example 3, page 117

$Z \sim N(0,1)$, Find $Z_{0.0125}$. That is

$$P(Z \geq z_{0.0125}) = 0.0125$$

check the table $\Rightarrow z_{0.0125} = 2.24$

What about $z_{0.05}$ and $z_{0.025}$?

$$\Rightarrow z_{0.05} = 1.645, \quad z_{0.025} = 1.960$$

Now we know how to compute $\Phi(y)$ by looking up the table for $Y \sim N(0,1)$.

Example 3, page 117

$Z \sim N(0,1)$, Find $Z_{0.0125}$. That is

$$P(Z \geq z_{0.0125}) = 0.0125$$

check the table $\Rightarrow z_{0.0125} = 2.24$

What about $z_{0.05}$ and $z_{0.025}$?

$$\Rightarrow z_{0.05} = 1.645, \quad z_{0.025} = 1.960$$

Now we know how to compute $\Phi(y)$ by looking up the table for $Y \sim N(0,1)$. What if Y is not standard normal?

Theorem 3.3-1

Theorem

If Y is $N(\mu, \sigma^2)$, then $X = \frac{Y-\mu}{\sigma}$ is $N(0,1)$

Proof: The idea is to show X has the same cdf as $N(0,1)$

$$P(X \leq x) = P(\frac{Y-\mu}{\sigma} \leq x) = P(Y \leq \sigma x + \mu) = \int_{-\infty}^{\sigma x+\mu} f(y)dy$$

$$= \int_{-\infty}^{\sigma x+\mu} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2}\right) dy$$

Theorem 3.3-1

coordinate change

$$w = \frac{y-\mu}{\sigma} \Longrightarrow \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}w^2)dw. \longrightarrow \text{cdf of } N(0,1).$$

Therefore, $P(X \leq x) = \Phi(x)$ and this completes the proof.

With the above theorem, for $X \sim N(\mu, \sigma^2)$

$$\begin{aligned} P(a \leq X \leq b) &= P(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

where $\Phi(\cdot)$ is the cdf of $N(0,1)$.

Example 4, page 118

$$X \sim N(3,16) \quad \text{Compute} \quad P(4 \le X \le 8), \quad P(0 \le X \le 5).$$

Example 4, page 118

$$X \sim N(3,16) \quad \text{Compute} \quad P(4 \le X \le 8), \quad P(0 \le X \le 5).$$

$$P(4 \le X \le 8) = P(\frac{4-3}{4} \le \frac{X-3}{4} \le \frac{8-3}{4})$$

$$= \Phi(1.25) - \Phi(0.25) = 0.8944 - 0.5987$$

$$P(0 \le X \le 5) = P(\frac{0-3}{4} \le \frac{X-3}{4} \le \frac{5-3}{4})$$

$$= \Phi(0.5) - \Phi(-0.75) = 0.6915 - 0.2266.$$

Relation between normal and χ^2 distribution

Theorem 3.3-2

If X is $N(\mu, \sigma^2)$ with $\sigma^2 > 0$, then

$$\frac{(X-\mu)^2}{\sigma^2} \sim \chi^2(1)$$

Proof: Let $V = \frac{(X-\mu)^2}{\sigma^2}$. Then consider the cdf of V,

$$G(v) = P(V \leq v) = P(-\sqrt{v} \leq Z \leq \sqrt{v})$$

where $Z = \frac{X-\mu}{\sigma}$, with $v \geq 0$.

Relation between normal and χ^2 distribution

$$G(v) = \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

Change the variable of integration $\quad z = \sqrt{y} \qquad \frac{dz}{dy} = \frac{1}{2\sqrt{y}}$

$$G(v) = 2\int_0^{v} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \frac{1}{2\sqrt{y}} dy = \int_0^{v} \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} dy, v \geq 0$$

$$g(v) = G'(v) = \frac{1}{\sqrt{2\pi}} v^{\frac{1}{2}-1} e^{-\frac{1}{2}v}, v \geq 0.$$

Now recall the pdf of $\chi^2(1)$:

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

Theorem 3.3-2

Since $g(v)$ is a pdf, then $\int_0^\infty g(v)dv = 1$

$$1 = \int_0^\infty \frac{1}{\sqrt{2\pi}} v^{\frac{1}{2}-1} e^{-\frac{1}{2}v} dv \xlongequal{x=\frac{1}{2}v} \frac{1}{\sqrt{\pi}} \int_0^\infty x^{\frac{1}{2}-1} e^{-x} dx$$

$$= \frac{1}{\sqrt{\pi}} \Gamma(\frac{1}{2}) \Rightarrow \Gamma(\frac{1}{2}) = \sqrt{\pi}$$

$$\Longrightarrow g(v) = \frac{1}{\Gamma(\frac{1}{2}) 2^{\frac{1}{2}}} v^{\frac{1}{2}-1} e^{-\frac{v}{2}}, v > 0$$