MATH 180A (Lecture A00)

mathweb.ucsd.edu/~ynemish/teaching/180a

Today: Gaussian (Normal) distribution
Normal approximation
Next: ASV 4.1

Week 6:

- Homework 4 due Friday, February 17

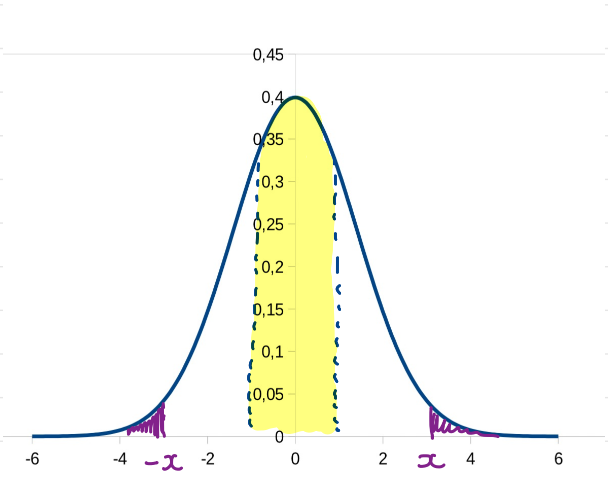
CDF of $\mathcal{N}(0,1)$

Suppose $X \sim \mathcal{N}(0,1)$. What is $P(|X| \le 1)$?

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

$P(-1 \le X \le 1)$

$$= \int_{-1}^{1} \varphi(t)\,dt = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-t^2/2}\,dt$$



Cannot use the polar coordinate trick.

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt \quad \text{- CDF of } X \sim \mathcal{N}(0,1)$$

- no simple explicit formula
- table of values of $\Phi(x)$ (for $x \ge 0$)

Normal table of values (Appendix E in textbook)

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	0.7704	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

This table gives $P(Z \leq z)$ where $Z \sim N(0,1)$, $z = x_i + y_j$

<u>Example</u> $\Phi(0.91) = P(Z \leq 0.91) = P(Z \leq 0.9 + 0.01) \approx 0.8186$

<u>Fact</u>: $\Phi(-x) = 1 - \Phi(x)$

$P(Z > 0.24) = 1 - P(Z \leq 0.24) = 1 - \Phi(0.24) = 1 - 0.5948 = 0.4052$

$P(-0.28 < Z < 0.59) = \Phi(0.59) - \Phi(-0.28) = \Phi(0.59) - (1 - \Phi(0.28))$

$= \Phi(0.59) + \Phi(0.28) - 1$

Normal table of values

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	0.7704	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

Exercise Let $Z \sim N(0,1)$

Find $x_0 \in \mathbb{R}$ such that $P(|Z| > x_0) \approx 0.704$

$$P(|Z| > x_0) = P(Z > x_0) + P(Z < -x_0) =$$

$$= 1 - \Phi(x_0) + (1 - \Phi(x_0)) = 2 \cdot (1 - \Phi(x_0)) \approx 0.704$$

$$1 - \Phi(x_0) \approx 0.352 \quad , \quad \Phi(x_0) \approx 1 - 0.352 = 0.648$$

$$x_0 \approx 0.38$$

Mean and variance of $X \sim N(0,1)$

$$E(X) = \int_{-\infty}^{+\infty} t\, f_X(t)\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t \cdot e^{-\frac{t^2}{2}}\, dt = \frac{-1}{\sqrt{2}}\, e^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} = 0$$

$$\frac{d}{dt}\left(-e^{-\frac{t^2}{2}}\right) = t e^{-\frac{t^2}{2}}$$

$$Var(X) = E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}}\, dt = \begin{cases} t e^{-\frac{t^2}{2}} = \left(-e^{-\frac{t^2}{2}}\right)' \\ \int u \cdot v' = u \cdot v - \int u' \cdot v \end{cases}$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}}\, t \cdot \left(-e^{-\frac{t^2}{2}}\right) \Big|_{-\infty}^{+\infty}}_{0} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} 1 \cdot \left(+e^{-\frac{t^2}{2}}\right) dt$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt = 1$$

General normal distribution $\quad N(\mu, \sigma^2)$

<u>Def</u> Let $\mu \in \mathbb{R}$ and $\sigma > 0$. Random variable X has normal (Gaussian) distribution with mean μ and variance σ^2 if the PDF of X is given by

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

We write $\quad X \sim N(\mu, \sigma^2)$

Using the density we can compute

$$E(X) = \mu \quad , \quad Var(X) = \sigma^2$$

"Gaussian distribution" = family of distributions

Relation between $X \sim N(\mu, \sigma)$ and $Z \sim N(0,1)$

Proposition Let $X \sim N(\mu, \sigma^2)$, $a \neq 0$, $b \in \mathbb{R}$.

Then the random variable $aX + b$ has normal distribution, $aX + b \sim N(a\mu + b, a^2\sigma^2)$

Using this proposition any Gaussian random variable can be written as a shifted and rescaled standart normal.

E.g., if $\sigma > 0$, $\mu \in \mathbb{R}$ and $Z \sim N(0,1)$, then

$$\sigma Z + \mu \sim N(\mu, \sigma^2)$$

If $X \sim N(\mu, \sigma^2)$, then $E(X) =$; $Var(X) =$

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

Example

Let $X \sim N(-3, 4)$

Find $P(X < 0.91)$; $P(X > 0.82)$; $P(-0.24 < X < 0.88)$

If $X \sim N(-3, 4)$, then $\frac{X+3}{2} \sim N(0,1)$, so

$$P(X < 0.91) = P\left(\frac{X+3}{2} < \frac{0.91+3}{2} \right) = P\left(\underbrace{\frac{X+3}{2}}_{\sim N(0,1)} < 1.955 \right)$$

$$= \Phi(1.955)$$

$$P(-0.24 < X < 0.88) = P\left(-\frac{0.24+3}{2} < \frac{X+3}{2} < \frac{0.88+3}{2} \right)$$

$$= \Phi(1.94) - \Phi(1.38)$$

The message:

If we have independent and identically distributed random variables $X_1, X_2, \dots, X_n$ with $E(X_1)=\mu$, $Var(X_1)=\sigma^2$, then for any $a<b$

$$\lim_{n\to\infty} P\left(a < \frac{X_1+X_2+\dots+X_n - \mu \cdot n}{\sqrt{n}\,\sigma} < b\right) = \int_a^b \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt$$

CENTRAL LIMIT THEOREM

Today: $X_1 \sim Ber(p)$; Last lecture: general case