Handout[1] 7

Nov 7, 2022

[Topics]:

Normal Distribution

1 Basic Introduction of Normal Distribution

Definition of Normal Distribution
A random variable X is normally distributed with mean μ and variance σ^2 ($-\infty < \mu < +\infty$ and $\sigma > 0$), denoted as $X \sim N(\mu, \sigma^2)$, if the *pdf* is given by:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in (-\infty, +\infty).$$

Graph Interpretation

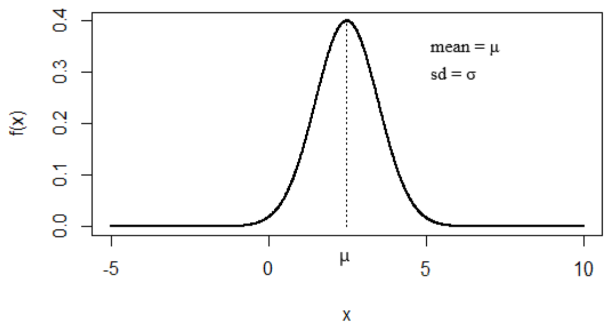

Figure 1: Normal Distribution with mean μ and variance σ^2

Some observations:
As shown in the graph above, $X \sim N(\mu, \sigma^2)$ has the largest $f(x)$ at $x = \mu$. It is symmetric to the line $x = \mu$. So a very straight fowrard property is, if x is closer to the mean μ, it is more likely that $X = x$ will happen.

Symmetry of Normal Distribution
As

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in (-\infty, +\infty).$$

We have

$$\begin{aligned} f(\mu + h \mid \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mu+h-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(h)^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mu-h-\mu)^2}{2\sigma^2}} \\ &= f(\mu - h \mid \mu, \sigma^2) \end{aligned}$$

[1]This handout is made by Hongkai Wang for Econ 329 Economic Statistics. It is adapted from Professor Wiseman's lectures, however, all errors are mine.

That is, that graph of $X \sim N(\mu, \sigma^2)$ is symmetry to its mean.

2 Linear Transformation of Normal Distribution

Suppose $X \sim N(\mu, \sigma^2)$, let $Y = aX + b$, where $a \neq 0$ and b are constant. Then $Y \sim N(a\mu + b, a^2\sigma^2)$.

The formal proof is by moment generating function $m.g.f$. However, there is a simple way to verify the mean and variance of Y. Recall that:

$$E(aX + b) = aE(X) + b, \text{ and } Var(aX + b) = a^2 Var(X).$$

Proof is on handout 6.

Definition of Standard Normal Distribution
Random variable X has a standard normal distribution, denoted as $X \sim N(0, 1)$ if its *pdf* is given by:

$$f(x \mid \mu = 0, \sigma^2 = 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in (-\infty, +\infty).$$

CDF of Standard Normal Distribution
For standard random variable X, the cdf is given by:

$$\Phi(x) = P(X \leq x) = \int_{-\infty}^{x} f(s \mid \mu = 0, \sigma^2 = 1) ds.$$

There is a chart of standard Normal distribution function (See appendix) that gives the value of $\Phi(x)$.

Standardization of Normal Distribution
For $X \sim N(\mu, \sigma^2)$, we have $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$. Now let $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$, then $Y = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \sim N(0, 1)$, which is standard normal. For standard normal, we have the chart and hence can determine the cdf $\Phi(y)$.
Verification
For $X \sim N(\mu, \sigma^2)$ and $Y = \frac{X - \mu}{\sigma}$,

$$\begin{aligned} E(Y) &= E(\frac{X - \mu}{\sigma}) \\ &= \frac{1}{\sigma} E(X - \mu) \\ &= \frac{1}{\sigma} \times 0 \\ &= 0 \end{aligned}$$

And

$$\begin{aligned} Var(Y) &= Var(\frac{X - \mu}{\sigma}) \\ &= (\frac{1}{\sigma})^2 Var(X) \\ &= (\frac{1}{\sigma})^2 \times \sigma^2 \\ &= 1 \end{aligned}$$

3 Linear Combinations of Normal Distribution

Normal plus Normal is still Normal. Normal minus Normal is still Normal.

Recall that, for random variables X and Y,

$$E(X + Y) = E(X) + E(Y).$$

This holds whenever X and Y are independent or not. (See handout 6).

And

$$Var(X+Y) = Var(X) + Var(Y).$$

This holds for X and Y are independent.

Also,

$$Var(aX) = a^2 Var(X).$$

See handout 6.

By the above equations, for X and Y independent, a and b constant, we have:

$$E(aX+bY) = aE(X) + bE(Y).$$

And

$$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y).$$

Linear Combination of Normal Distribution
For independent normal random variables $X_1, X_2, \cdots, X_n$, where $X_i \sim N(\mu_i, \sigma_i^2)$. The sum $Y = \sum_{i=1}^n X_i$ is noraml with mean to be $\sum_{i=1}^n \mu_i$ and variance to be $\sum_{i=1}^n \sigma_i^2$.
Verification

$$\begin{aligned} E(Y) &= E(\sum_{i=1}^n X_i) \\ &= \sum_{i=1}^n E(X_i) \\ &= \sum_{i=1}^n \mu_i \end{aligned}$$

And

$$\begin{aligned} Var(Y) &= Var(\sum_{i=1}^n X_i) \\ &= \sum_{i=1}^n Var(X_i) \\ &= \sum_{i=1}^n \sigma_i^2 \end{aligned}$$

General Case
For independent normal random variables $X_1, X_2, \cdots, X_n$, where $X_i \sim N(\mu_i, \sigma_i^2)$. Then $Y = \sum_{i=1}^n a_i X_i$ is noraml with mean to be $\sum_{i=1}^n a_i\mu_i$ and variance to be $\sum_{i=1}^n a_i^2\sigma_i^2$.

4 Sample Mean of Normal Distribution

For independent and identically distributed(i.i.d) random variables $X_1, X_2, \cdots, X_n$ with $X_i \sim N(\mu, \sigma^2)$, the sample mean $Y = \frac{1}{n}\sum_{i=1}^n X_i$ is normal and $Y \sim N(\mu, \frac{\sigma^2}{n})$.

Verification

$$
\begin{aligned}
E(Y) &= E(\frac{1}{n}\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\
&= \frac{1}{n}\cdot n \cdot \mu \\
&= \mu
\end{aligned}
$$

And

$$
\begin{aligned}
Var(Y) &= Var(\frac{1}{n}\sum_{i=1}^{n} X_i) \\
&= (\frac{1}{n})^2 \sum_{i=1}^{n} Var(X_i) \\
&= (\frac{1}{n})^2 \cdot n \cdot Var(X_i) \\
&= \frac{\sigma^2}{n}
\end{aligned}
$$

5 Exercises

1. Basis of Normal
Suppose random variable X has *pdf* :

$$f(x) = \frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-7)^2}{32}}, \quad x \in (-\infty, +\infty).$$

What is mean and variance of X?
Solution By comparing the form above and $f(x \mid \mu, \sigma^2)$, we have mean $\mu = 7$ and variance $\sigma^2 = 16$.

2. Symmetry of Normal
Suppose random variable $X \sim N(5, 25)$, if known $f(5+\delta) = 0.067$, what is $f(5-\delta)$?
Solution By symmetry of normal, we have $f(5-\delta) = f(5+\delta) = 0.067$.

3. Symmetry of Normal
Suppose random variable $X \sim N(0, 1)$, if known $f(a)$, what is $f(-a)$?
Solution By symmetry of normal, we have $f(-a) = f(a)$.

4. Symmetry of Normal
Suppose random variable $X \sim N(0, 1)$, if known $\Phi(a)$, what is $\Phi(-a)$?
Solution By symmetry of normal, $P(X \leq -a) = P(X \geq a)$.

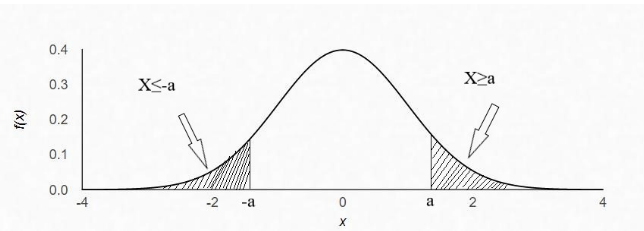

Figure 2: Normal Distribution with $\mu = 0$ and $\sigma^2 = 1$

As shown in the graph, the area represents $P(X \leq -a)$ and the area of $P(X \geq a)$ are identical. Hence $P(X \leq -a) = P(X \geq a)$, then $\Phi(-a) = 1 - \Phi(a)$.

5. Linear Transformation
Suppose $X \sim N(1314, 520)$. What is the distribution of $Y = \frac{\sqrt{1314}}{\sqrt{520}}X - \frac{1314\times\sqrt{1314}-520\times\sqrt{520}}{\sqrt{520}}$?

Solution Y is normal. So we need to specify the parameters. The mean of Y is :

$$\begin{aligned}
E(Y) &= E(\frac{\sqrt{1314}}{\sqrt{520}}X - \frac{1314\times\sqrt{1314}-520\times\sqrt{520}}{\sqrt{520}}) \\
&= \frac{\sqrt{1314}}{\sqrt{520}}E(X) - \frac{1314\times\sqrt{1314}-520\times\sqrt{520}}{\sqrt{520}} \\
&= \frac{\sqrt{1314}}{\sqrt{520}} \times 1314 - \frac{1314\times\sqrt{1314}-520\times\sqrt{520}}{\sqrt{520}} \\
&= 520
\end{aligned}$$

And

$$\begin{aligned}
Var(Y) &= Var(\frac{\sqrt{1314}}{\sqrt{520}}X - \frac{1314\times\sqrt{1314}-520\times\sqrt{520}}{\sqrt{520}}) \\
&= (\frac{\sqrt{1314}}{\sqrt{520}})^2 Var(X) \\
&= \frac{1314}{520} \times 520 \\
&= 1314
\end{aligned}$$

Therefor, $Y \sim N(520, 1314)$.

6. Standardization
Suppose $X \sim N(4, 16)$, what is $P(X \leq 8)$?

Solution For X not a standard normal, we need to make it Standard:

$$Y = \frac{X-\mu}{\sigma} = \frac{X-4}{4} \sim N(0, 1).$$

So

$$\begin{aligned}
P(X \leq 7) &= P(\frac{X-4}{4} \leq \frac{7-4}{4}) \\
&= P(Y \leq \frac{3}{4}) \\
&= \Phi(\frac{3}{4})
\end{aligned}$$

As shown in the chart, $\Phi(\frac{3}{4}) = 0.7734$.

7. Standardization
Suppose $X \sim N(4, 16)$, what is $P(X \geq 8)$?

Solution

$$\begin{aligned}
P(X \geq 8) &= P(\frac{X-4}{4} \leq \frac{8-4}{4}) \\
&= P(Y \geq 1) \\
&= 1 - \Phi(1)
\end{aligned}$$

As shown in the chart, $\Phi(1) = 0.8413$, then $1 - \Phi(1) = 0.1587$.

8. Linear Transformation
Suppose $X \sim N(4, 16)$. Find a transformation of X to be noraml with mean 7 and variance 9.

Solution Suppose $Y = aX + b \sim N(7,9)$, where $X \sim N(4,16)$.
Then we have $E(Y) = E(aX+b) = aE(X) + b = 4a + b$ and $Var(Y) = Var(aX + b) = a^2Var(X) = a^2 \times 16$.
Now set $4a + b = 7$ and $16a^2 = 9$, we have $a = \frac{3}{4}$ and $b = 4$.
i,e. $Y = \frac{3}{4}X + 4$ transforms X to $N(7,9)$.

9. Linear Combination
Suppose $Var(X) = 8$, what is $Var(-X)$?
Solution $Var(-X) = (-1)^2Var(X) = Var(X) = 8$.

10. Linear Combination
Suppose $Var(X) = 8$ and $Var(Y) = 7$, X and Y are independent. What is $Var(X+Y)$?
Solution $Var(X+Y) = Var(X) + Var(Y) = 15$.

11. Linear Combination
Suppose $Var(X) = 8$ and $Var(Y) = 7$, X and Y are independent. What is $Var(X-Y)$?
Solution $Var(X-Y) = Var(X + (-1)Y) = Var(X) + (-1)^2Var(Y) = 15$.

12. Linear Combination
Suppose $X \sim N(3,8)$, $Y \sim N(4,9)$, X and Y are independent. What is the distribution of $X+Y$?
Solution Normal plus normal is still normal. So we specify the parameters(mean and variance).
$E(X+Y) = E(X) + E(Y) = 7$, $Var(X+Y) = Var(X) + Var(Y) = 17$. So $X + Y \sim N(7,17)$.

13. Linear Combination
Suppose $X \sim N(3,8)$, $Y \sim N(4,9)$, X and Y are independent. What is the distribution of $X-Y$?
Solution Normal minus normal is still normal. So we specify the parameters(mean and variance).
$E(X-Y) = E(X) - E(Y) = -1$, $Var(X-Y) = Var(X) + Var(Y) = 17$. So $X + Y \sim N(-1,17)$.

14. Sample Mean
Suppose a random sample of 32 observations is taken from the normal distribution with mean 6 and variance 64. What is the distribution of the sample mean?
Solution Let $Y = \frac{1}{32}\sum_{i=1}^{32} X_i$ denotes the sample mean, where $X_i \sim N(6,64)$.
Then Y is still normal. $E(Y) = E(\frac{1}{32}\sum_{i=1}^{32} X_i) = \frac{1}{32}\sum_{i=1}^{32} E(X_i) = \frac{1}{32} \times 32 \times 6 = 6$. And $Var(Y) = Var(\frac{1}{32}\sum_{i=1}^{32} X_i) = (\frac{1}{32})^2 \sum_{i=1}^{32} Var(X_i) = (\frac{1}{32})^2 \times 32 \times 64 = 2$. So the sample mean is noraml $N(6,2)$.

15. Sample Mean
Suppose a random sample of 32 observations is taken from the normal distribution with mean 6 and variance 64. What is the probability that the sample mean is in the interval [5, 7]?
Solution As in 14, sample mean $Y = \frac{1}{32}\sum_{i=1}^{32} X_i \sim N(6,2)$.
So

$$\begin{aligned}
P(5 \le Y \le 7) &= P(\frac{5-6}{\sqrt{2}} \le \frac{Y-6}{\sqrt{2}} \le \frac{7-6}{\sqrt{2}}) \\
&= P(-\frac{1}{\sqrt{2}} \le Z \le \frac{1}{\sqrt{2}}),\ where\ Z \sim N(0,1) \\
&= \Phi(\frac{1}{\sqrt{2}}) - \Phi(-\frac{1}{\sqrt{2}}) \\
&= 2\Phi(\frac{1}{\sqrt{2}}) - 1
\end{aligned}$$