4 THE NORMAL DISTRIBUTION

4.1 Introduction

A very important continuous distribution. Let the random variable $X$ have a normal distribution with mean $\mu$ and variance $\sigma^2$. We write $X \sim N(\mu, \sigma^2)$. Note that the variance is given in this description.

Let $Y = aX + b$. Then

$$E(Y) = a\mu + b, \quad \text{Var}(Y) = a^2\sigma^2,$$

and $Y$ also has a normal distribution, $Y \sim N(a\mu + b, a^2\sigma^2)$. In particular, if

$$Z = \frac{X - \mu}{\sigma} \quad (a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma})$$

then $E(Z) = 0$ and $\text{Var}(Z) = 1$, so $Z \sim N(0, 1)$. We call $Z$ the standard normal distribution.

If $Z \sim N(0, 1)$ and $X = aZ + b$, then $X \sim N(a\mu + b, a^2)$. 

Probability density functions of $N(\mu, \sigma^2)$ for different means and variances.
Tables list values for the standard normal distribution. Note: Exam tables will be the same as that in the text book, and a copy is available on line.
For $X \sim N(\mu, \sigma^2)$ there are usually two types of problems:

(i) Given an interval for $X$ find the probability;

(ii) Given a probability, find the value(s) of $X$ (the inverse problem).

**Example 1. Let $Z \sim N(0, 1)$. find**

(i) $P(Z < 1.0)$,

(ii) $P(Z < -1.0)$,

(iii) $P(Z > 1.0)$,

(iv) $P(Z < 2.51)$,

(v) $P(|Z| < 1.96)$,

(vi) $z$ such that $P(Z < z) = 0.9505$,

(vii) $z$ such that $P(|Z| < z) = 0.95$,

(viii) $P(-1.5 < Z < 2.7)$.

**Solution**
Example 2. Let $X \sim N(5, 16)$. Find

(i) $P(X < 0)$,

(ii) $P(X > 10)$,

(iii) $P(-5 \leq X \leq 7)$,

(iv) $c$ such that $P(|X - \mu| < c) = 0.95$.

Solution
Example 3. Let $X \sim N(\mu, 0.25)$, and suppose $P(X < 5.1) = 0.9772$. Find $\mu$.

Solution
4.2 SUM OF NORMAL RANDOM VARIABLES

Let $X \sim N(\mu_X, \sigma^2_X)$, $Y \sim N(\mu_Y, \sigma^2_Y)$, and put $W = X \pm Y$. Then $W \sim N(\mu_W, \sigma^2_W)$, where

\[
\mu_W = \mu_X \pm \mu_Y \\
\text{and } \sigma^2_W = \text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y) \pm 2 \text{Cov}(X,Y).
\]

If $X$ and $Y$ are independent, then $\text{Cov}(X,Y) = 0$, so

\[
\mu_W = \mu_X \pm \mu_Y \\
\text{and } \sigma^2_W = \sigma^2_X + \sigma^2_Y.
\]

This result can be extended to a sum of several normal random variables.

**Result**

Let $X_1, X_2, \cdots, X_n$ be independent normal random variables with mean $\mathbb{E}(X_i) = \mu_i$ and $\text{Var}(X_i) = \sigma^2_i$, $i = 1, 2, \cdots, n$. Put

\[Y = \sum_{i=1}^{n} X_i.\]

Then $Y \sim N(\mu_Y, \sigma^2_Y)$, where

\[
\mu_Y = \mathbb{E} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \sum_{i=1}^{n} \mu_i
\]

and

\[
\sigma^2_Y = \text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i) \text{ (Independence)} = \sum_{i=1}^{n} \sigma^2_i.
\]
Example 4. Suppose the income of executives is normally distributed with the incomes of men and women independent of each other. The means and standard deviations (in $10,000) respectively are 20, 2.5 for men and 15, 2.0 for women.

(a) What is the probability that a randomly chosen female executive earns more than a randomly chosen male executive?

(b) Find the probability that an executive couple has a combined income of more than $400,000.

Solution
Example 5. A machine makes washers with hole diameters that are normally distributed, with mean 15.2 mm and variance 0.03 mm$^2$. Another machine makes bolts with diameters that are normally distributed, with mean 15.0 mm and variance 0.01 mm$^2$.

(a) What is the probability that a randomly selected bolt with fit through a randomly selected washer?

(b) What should be the mean diameter of the washer holes if 99% of the bolts are to fit the washers?

Solution
4.3 Normal approximation to the Binomial Distribution

Let $X \sim \text{Bin}(n, p)$. Then $\mu_X = np$, $\sigma_X = \sqrt{np(1-p)} = \sqrt{npq}$, where $q = 1 - p$. Let $Y \sim N(np, np(1-p))$. As $n$ increases, the distribution of $X$ is more closely approximated by that of $Y$. (See the plots of the pmf and pdf in the textbook.) Thus

$$\frac{X - np}{\sqrt{np(1-p)}} \sim N(0, 1).$$

The approximation is good provided:

1. $n \geq 30$,
2. $np > 5$
3. $n(1-p) > 5$.

**Continuity correction** further improves the approximation.

**Example 6.** $X \sim \text{Bin}(100, 0.4)$. Find approximately

(a) $P(X \leq 30)$,
(b) $P(X > 50)$,
(c) $P(X \geq 50)$,
(d) $P(30 \leq X \leq 50)$,
(e) $P(X = 50)$.

**Solution**
Exercise
Find the exact values for the above example from Excel.
[(a) 0.0248, (b) 0.0168, (c) 0.0281, (d) 0.932-0.0148=0.9684, (e) 0.0103]

**Continuity Correction Rule:**
\[ P(X \leq a) \approx P \left( Z \leq \frac{a+0.5-\mu}{\sigma} \right) \]
We can always write the required probability as \( P(X \leq a) \).

Exercise

©Nazim Khan, 2007