The University of Sydney

MATH1005 Statistics

maths.usyd.edu.au/u/UG/WS/WS1005/

Summer/Winter/Semester2 **Tutorial Solutions 7** 2015

This tutorial explores continuous distributions, with a special focus on the Normal.
Most questions involve hand calculations.
Draw a sketch of the appropriate Normal before looking up the Normal table or using R.

For a continuous random variable	
probability density function (pdf)	$f(x)\ \forall x$ (often represented as graph)
	$f(x) = F'(x)$
cumulative distribution function (CDF)	$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)dy$
	$P(X \le x) = P(X < x)$ as $P(X = x) = 0\ \forall x$
expected value or mean	$E(X) = \int_x xf(x)dx$
expected value of function	$E(g(X)) = \int_x g(x)f(x)dx$
Eg	$E(X^2) = \int_x x^2 f(x)dx$
variance	$Var(X) = E(X^2) - (E(X))^2$

General Normal Distribution

A General Normal random variable $X \sim N(\mu, \sigma^2)$ has mean μ and variance σ^2.

The pdf is $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Standardising a Normal

If $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0, 1)$, then

$$P(X \le x) = P\Big(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\Big) = P\Big(Z \le \frac{x-\mu}{\sigma}\Big)$$

1. Continuous Distribution by hand

A random variable X has the following CDF:

$$F(x) = \begin{cases} 0 & x \le 0 \\ x/2 & x \in (0, 2) \\ 1 & x \ge 2 \end{cases}$$

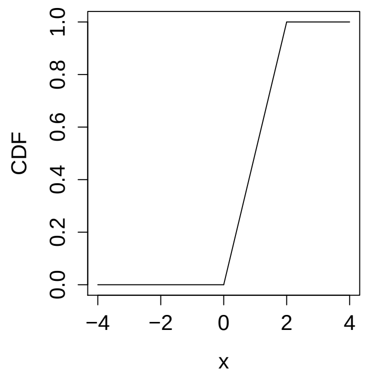
(a) Sketch the CDF.

(b) Find the pdf and sketch it.

(c) Calculate $E(X)$ and $Var(X)$. Does this match up with your sketches?

Solution

(a) Check your hand drawn sketch against this R output:



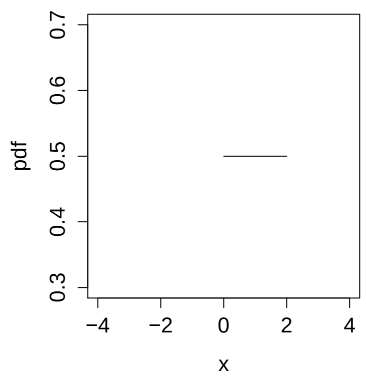
(b) To get the pdf, we differentiate the CDF. Hence,

$$f(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{2} & x \in (0,2) \\ 0 & x \ge 2 \end{cases}$$

giving

$$f(x) = \begin{cases} \dfrac{1}{2} & x \in (0,2) \\ 0 & \text{otherwise} \end{cases}$$

Again, check your hand drawn sketch against this R output:



(c) $E(X) = \displaystyle\int_{-\infty}^{\infty} xf(x)dx = \frac{1}{2}\int_0^2 xdx = \frac{1}{2}[\frac{x^2}{2}]_0^2 = 1$

Comment: As the pdf is symmetric, we expect the mean at $x = 1$.

$E(X^2) = \displaystyle\int_{-\infty}^{\infty} x^2 f(x)dx = \frac{1}{2}\int_0^2 x^2 dx = \frac{1}{2}[\frac{x^3}{3}]_0^2 = 8/6$

Hence $Var(X) = 8/6 - (1)^2 = 1/3$, and $SD(X) \approx 0.58$.

Comment: Looking at the pdf, it seems reasonable that the standard deviation would be 0.58. (The mean is 1 and the full range of data is only 2).

2. Standard Normal Probabilities

Given $Z \sim N(0, 1)$.

(a) By looking up the Normal tables, find $P(Z \leq 1)$ and $P(Z \leq -1)$.

(b) Fill out the following table. What do these probabilities represent?

$P(-1 \leq Z \leq 1)$	$P(-2 \leq Z \leq 2)$	$P(-3 \leq Z \leq 3)$
0.68		

(c) Find the values of a, b and c such that $P(Z \leq a) = 0.7291, P(Z > b) = 0.10$ and $P(|Z| < c) = 0.90$.

(d) Use R to confirm your answers.

```
> pnorm(1)          % This calculates P(Z <=1)
> qnorm(.7291)      % This calculates a    where  P(Z <=a ) = 0.7291
```

Solution

(a) From Normal tables, $P(Z \leq 1) = 0.8413$
$P(Z \leq -1) = 1 - P(Z \leq 1) = 0.1587$.

(b)

$P(-1 \leq Z \leq 1)$	$P(-2 \leq Z \leq 2)$	$P(-3 \leq Z \leq 3)$
0.6827 or 0.68	0.9545 or 0.95	0.9973 or 0.99

Comment: This table represents the probability of being 1,2 and 3 standard deviations away from the mean. Eg The probability of being within 1 standard deviation of the mean is approximately 68%.

(c) Using the Normal table,
if $P(Z \leq a) = 0.7291$, then $a = 0.6101$
if $P(Z > b) = 0.10$, then $P(Z < b) = 0.90$, so $b = 1.2816$
if $P(|Z| < c) = 0.90$, then $P(Z < c) = 0.95$, hence $c = 1.645$

(d) R commands given.

3. Normal Probabilities

Given $X \sim N(5, 4)$.

(a) What is $P(X = 1)$?

(b) Calculate $P(X \leq 1)$ and $P(-2 \leq X \leq 1)$.

(c) Use R to confirm your answers.

```
> pnorm(1,5,2)       % This calculates P(X <=1)
```

Solution

(a) For any continuous distribution, $P(X = x) = 0$ for all x.

(b) $P(X \leq 1) = P(\frac{X-5}{2} \leq \frac{1-5}{2}) = P(Z \leq -2) = 1 - P(Z \leq 2) = 1 - \Phi(2) = 0.0228$
$P(-2 \leq X \leq 1) = P(\frac{-2-5}{2} \leq \frac{X-5}{2} \leq \frac{1-5}{2}) = P(-3.5 \leq Z \leq -2) = P(2 \leq Z \leq 3.5) = \Phi(3.5) - \Phi(2) = 0.0226$.
(c)

```
> pnorm(1,5,2)          OR > pnorm(-2)
[1] 0.02275013
```

```
> pnorm(3.5)-pnorm(2)
[1] 0.0225175
```

4. Standardising Normal Distribution

X is the number of hours per week spent watching TV by a young child. X is approximately normal with mean $\mu = 15$ hours and standard deviation $\sigma = 2$ hours. For a child chosen at random, find the probability that the child watches between 15 and 19 hours of TV per week.

Solution

X = the number of hours per week spent watching TV by a young child $\sim N(\mu = 15, \sigma^2 = 4)$.

$P(15 < X < 19) = P(\frac{15-15}{2} < \frac{X-15}{2} < \frac{19-15}{2}) = P(0 < Z < 2) = \Phi(2) - \Phi(0) = \Phi(2) - 0.5 = 0.9772 - 0.5 = 0.4772.$

Check in R:

```
> pnorm(19,15,2) - pnorm(15,15,2) OR  > pnorm(2)-pnorm(0)
[1] 0.4772499
```

5. Expectation of Normal Distribution

Suppose that X is the breaking strength of a rope in pounds, with distribution $N(100, 16)$. Each 30 metre coil of rope brings a return of \$25 provided $X > 95$. If $X \le 95$ then the rope is used for a different purpose and a return of only \$10 per coil is made. Find the expected return per coil.

Solution

X = the breaking strength of a rope in pounds $\sim N(100, 16)$.

(1) Good quality rope : If $P(X > 95)$ then \$25 return.

$P(X > 95) = P(\frac{X-100}{4} > \frac{95-100}{4}) = P(Z > -5/4) = P(Z < 5/4) = \Phi(5/4) = 0.8943.$

(2) Poor quality rope : If $P(X < 95)$ then \$10 return.

$P(X < 95) = 1 - P(X > 95) = 0.1057.$

Hence expected return per coil of rope is:
$E(X) = P(X > 95) \times \$25 + P(X < 95) \times \$10 = 0.8943 \times \$25 + 0.1057 \times \$10 = \$23.4145.$

6. Normal Distribution

Each day random samples of a city's water supply are analysed and the concentration of fluoride (parts per million, ppm) is measured. It is found that the distribution of this concentration is approximately normal with mean of 1.3 ppm and a standard deviation of 0.15 ppm.

(a) On what proportion of days is the concentration between 1.45 and 1.53 ppm?

(b) What is the 80th percentile of the distribution? What does this represent?

Solution

X = the concentration of fluoride $\sim N(1.3, 0.15^2)$.

(a) $P(1.45 \le X \le 1.53) = P(\frac{1.45-1.3}{0.15} \le \frac{X-1.3}{0.15} \le \frac{1.53-1.3}{0.15}) = P(1 < Z < 1.53) = \Phi(1.53) - \Phi(1) = 0.9370 - 0.8413 = 0.0957$

Check in R:

```
> pnorm(1.53,1.3,0.15)-pnorm(1.45,1.3,0.15)
[1] 0.09605838
```

(b) We want to find q such that $P(X < q) = 0.8$.

$$P(X < q) = P\left(\frac{X - 1.3}{0.15} < \frac{q - 1.3}{0.15}\right) = P\left(Z < \frac{q - 1.3}{0.15}\right) = 0.8.$$

From looking up Normal tables, we find that $\dfrac{q - 1.3}{0.15} = 0.8416$. So $q = 1.4262$.

Represents: We expect the level of fluoride to be less than 1.4262 80% of the time.

Check in R:

```
> pnorm(1.4262,1.3,0.15)
[1] 0.7999194     # which is approximately 80%
```