This document describes how to use R to calculate probabilities associated with common distributions. R has a number of built in functions for calculations involving probability distributions, both discrete and continuous, including the binomial, normal, Poisson, geometric, gamma, beta, and others we have seen.

For each distribution, R has four primary functions. Each function has a one letter prefix followed by the root name of the function. The names make mnemonic sense for continuous random variables but are used in both cases. For example `dnorm` is the height of the *density* of a normal curve while `dbinom` returns the *probability* of an outcome of a binomial distribution. Here is a table of these commands.

	Meaning	
Prefix	Continuous	Discrete
d	density	probability (pmf)
p	probability (cdf)	probability (cdf)
q	quantile	quantile
r	random	random

Distribution	Root
Binomial	`binom`
Geometric	`geom`
Poisson	`pois`
Normal	`norm`
Gamma	`gamma`
Beta	`beta`
t	`t`
F	`F`
Chi-square	`chisq`

The Binomial Distribution.

The binomial distribution is applicable for counting the number of outcomes of a given type from a prespecified number n independent trials, each with two possible outcomes, and the same probability of the outcome of interest, θ. The distribution is completely determined by n and θ. The probability mass function is defined as:

$$\Pr\{X = k\} = \binom{n}{k}\theta^k(1-\theta)^{n-k}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

is called a binomial coefficient.

Probability Calculations.— In R the function `dbinom` returns binomial probabilities. There are three required arguments: the value(s) for which to compute the probability (k), the number of trials (n), and the success probability for each trial (θ). Usually, these are specified in this order.

For example, here we find the complete distribution when $n = 5$ and $\theta = 0.1$.

```
> dbinom(0:5, 5, 0.1)

[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001
```

Here `0:5` means the sequence 0, 1, 2, 3, 4, 5 to R.

The arguments can also be specified in arbitrary order by providing names of the arguments.

```
> dbinom(size = 5, prob = 0.1, x = 0:5)

[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001
```

If we want to find the single probability of exactly 10 successes in 100 trials with $\theta = 0.1$, we do this.

```
> dbinom(10, 100, 0.1)

[1] 0.1318653
```

The function `pbinom` is useful for summing consecutive binomial probabilities. With $n = 5$ and $\theta = 0.1$, here are some example calcuations.

$$
\begin{aligned}
\Pr\{X \le 2\} &= \text{pbinom}(2,5,0.1) \doteq 0.99144 \\
\Pr\{X \ge 3\} &= 1 - \Pr\{X \le 2\} = 1 - \text{pbinom}(2,5,0.1) \doteq 0.00856 \\
\Pr\{1 \le X \le 3\} &= \Pr\{X \le 3\} - \Pr\{X \le 0\} = \text{pbinom}(3,5,0.1) - \text{pbinom}(0,5,0.1) \doteq 0.40905
\end{aligned}
$$

An alternative is to use `dbinom` in conjunction with `sum`.

$$
\begin{aligned}
\Pr\{X \le 2\} &= \texttt{sum}(\texttt{dbinom}(0:2, 5, 0.1)) \doteq 0.99144 \\
\Pr\{X \ge 3\} &= \texttt{sum}(\texttt{dbinom}(3:5, 5, 0.1)) \doteq 0.00856 \\
\Pr\{1 \le X \le 3\} &= \texttt{sum}(\texttt{dbinom}(1:3, 5, 0.1)) \doteq 0.40905
\end{aligned}
$$

Quantiles.— We can also find the quantiles of a binomial distribution. For example, here is the 90th percentile of a binomial distribution with $n = 200$ and $\theta = 0.3$. The function `qbinom` finds the quantile.

```
> qbinom(0.9, 200, 0.3)

[1] 68
```

Random variable generation.— The last function for the binomial distribution is used to take random samples. Here is a random sample of 20 binomial random variables drawn from the binomial distribution with $n = 10$ and $\theta = 0.5$.

```
> rbinom(20, 10, 0.5)

 [1] 4 5 6 9 6 6 4 6 6 3 6 4 6 5 5 3 6 6 4 4
```

Normal Distribution

Normal distributions have symmetric, bell-shaped density curves that are described by two parameters: the mean μ and the standard deviation σ. The two points of a normal density curve that are the steepest—at the "shoulders" of the curve— are precisely one standard deviation above and below the mean.

Heights of individual corn plants may be modeled as normally distributed with a mean of 145 cm and a standard deviation of 22 cm (Samuels and Witmer, third edition, exercise 4.29). Here are several example normal calculations using R.

Probabilities.— Find the proportion of plants:
. . . larger than 100cm;

```
> 1 - pnorm(100, 145, 22)

[1] 0.979595
```

. . . between 120cm and 150cm:

```
> pnorm(150, 145, 22) - pnorm(120, 145, 22)

[1] 0.461992
```

. . . 150cm or less:

```
> pnorm(150, 145, 22)

[1] 0.5898942
```

Qunatiles.— Find the 75th percentile.

```
> qnorm(0.75, 145, 22)

[1] 159.8388
```

Find the endpoints of middle 95% of the distribution.

```
> qnorm(c(0.025, 0.975), 145, 22)

[1] 101.8808 188.1192
```

Other Distributions

Other distributions work in a similar way. Details on how to express the parameters for different probability distributions can be found from the help files. For example, to learn about find Poisson probabilities, type `?dpois`.