Basic Simulation (Gaussian)

Suppose that X is a random variable with PDF $\pi(x)$. We are interested in calculating

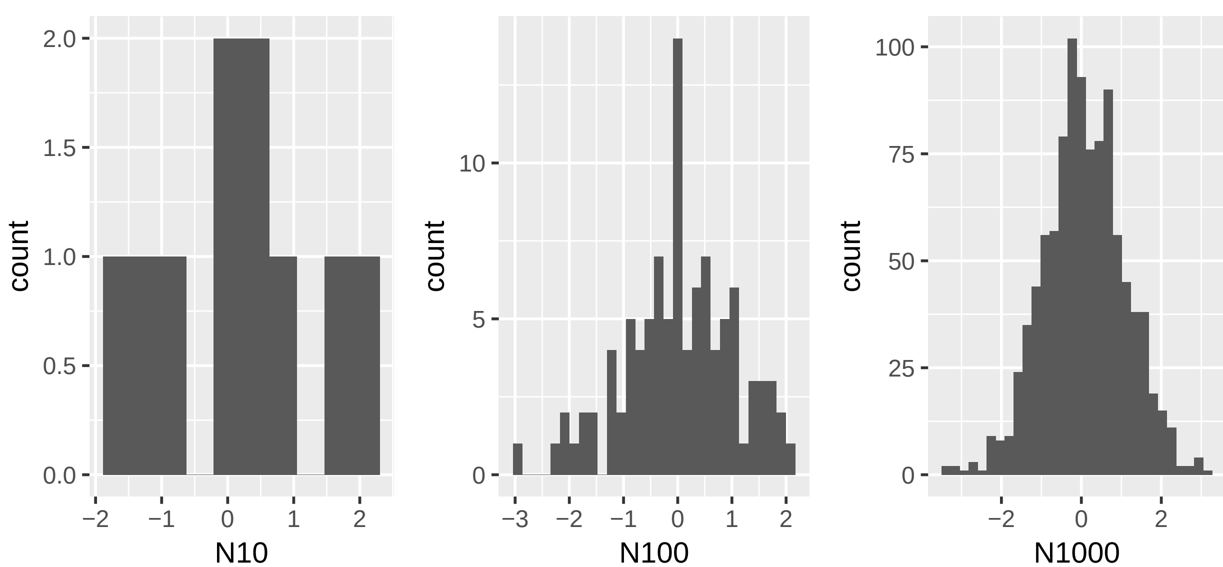
$$\mathbb{E}_\pi(h(X)) = \int h(x)\pi(x)\mathrm{dx}$$

This might be an interval difficult to integrate and might not have a closed-form solution, so we want to generate arbitrary points from the distribution $\pi(x)$ to approximate this integral. There are cases where we only know $\pi(x)$ up to a multiplicative/normalizing constant. In these cases, importance sampling is used, which we will cover later.

rnorm

Suppose we know that $X \sim$ Gaussian(0,1). We draw randomly n times from the standard Normal distribution using the function `rnorm`.

```
set.seed(21)
# 10 random draws from standard normal
N10 <- rnorm(n = 10, mean = 0, sd = 1)
# 100 random draws from standard normal
N100 <- rnorm(n = 100, mean = 0, sd = 1)
# 1000 random draws from standard normal
N1000 <- rnorm(n = 1000, mean = 0, sd = 1)
# plot histogram of x_sim, follows a Gaussian(0,1)
n10 <- qplot(N10, geom="histogram", bins = 10)
n100 <- qplot(N100, geom="histogram", bins = 30)
n1000 <- qplot(N1000, geom="histogram", bins = 30)
grid.arrange(n10, n100, n1000, ncol = 3)
```

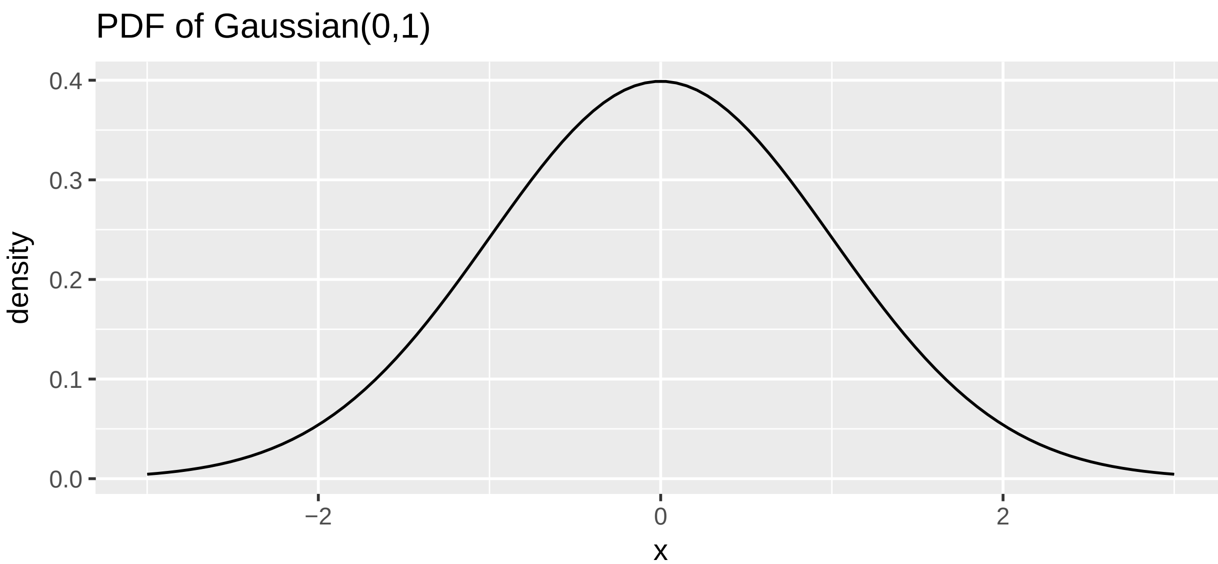


dnorm

`dnorm` returns the probability density for the Gaussian distribution given parameters `x`, μ, and σ, where Gaussian density function is

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(x-\mu)^2}{2\sigma^2}\Big\}$$

```
# create a vector of z-scores
x_seq <- seq(-3,3,length.out = 100)
# find normal densities at each x
f_x_seq <- dnorm(x_seq)
# plot
qplot(x = x_seq, y = f_x_seq, geom = "line",
      xlab = "x", ylab = "density", main = "PDF of Gaussian(0,1)")
```

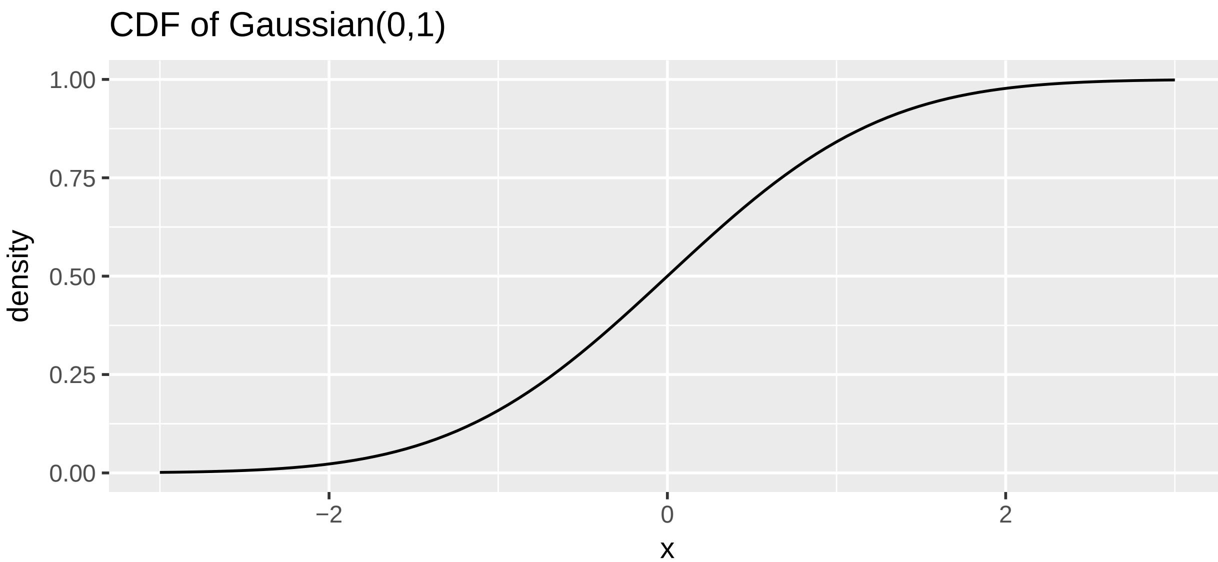


pnorm

`pnorm` returns the value of $F(x) = \mathbb{P}(X \leq x)$ at each \$x, which is equivalent to

$$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(x-\mu)^2}{2\sigma^2}\Big\}dx$$

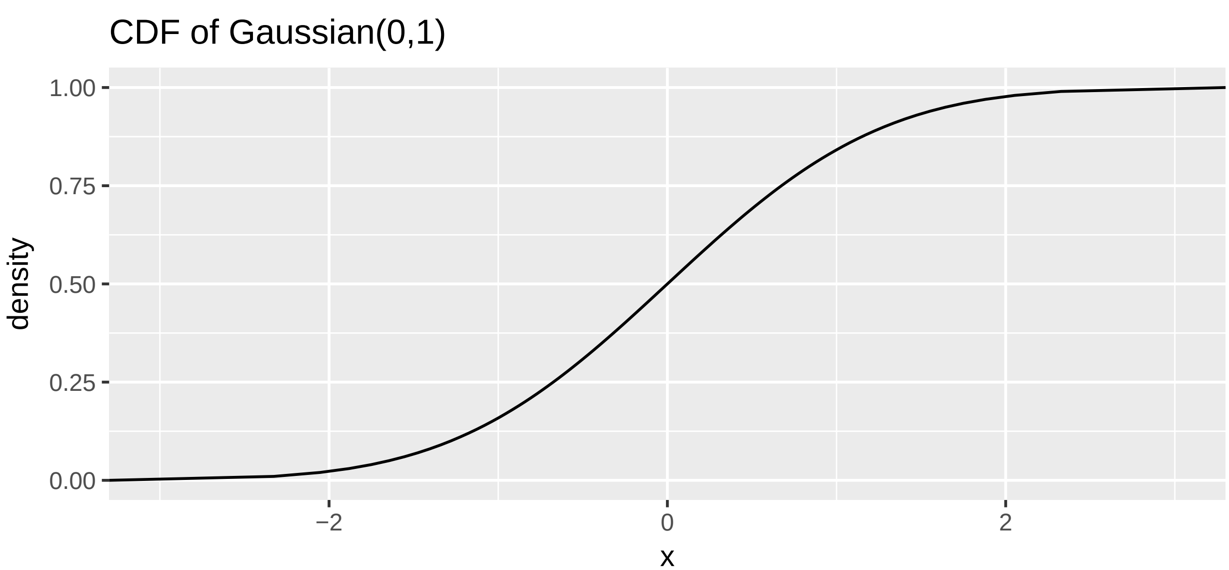
```
pvalues <- pnorm(x_seq)
plt1 = qplot(x = x_seq, y = pvalues, geom = "line",
             xlab = "x", ylab = "density", main = "CDF of Gaussian(0,1)") +
  xlim(c(-3,3))
plt1
```



qnorm

`qnorm` is inverse of `pnorm`. It returns x from a given $F(x)$.

```
quantiles <- seq(0, 1, by = .01)
qvalues <- qnorm(quantiles)
plt2 <- qplot(x = qvalues, y = quantiles, geom = "line",
              xlab = "x", ylab = "density", main = "CDF of Gaussian(0,1)") +
  xlim(c(-3,3))
plt2
```



Lastly, consider the function $h(x) = 10e^{-2\|x-5\|}$. We want to find $\mathbb{E}(h(X))$.

```
h <- function(x) 10*exp(-2*abs(x-5))
# calculate h(x) at each of the random points drawn from standard normal and then approximate E(h(X))
mean(h(N1000))
```

```
## [1] 0.003719365
```