More Probability, Continous Random Variables

Rob McCulloch

1. Continuous Random Variables
2. Expection, Mean, Variance, Covariance

1. Continuous Random Variables

Sometimes it is inconvenient to list out all the possible values a random variable can take on.

For example, we don't want to list all the possible times a patient could live for.

In this case we let our random variables take on on value in R, or any value in a subset of R.

For example we might think of the time our patient live to be any value in the subset of R given by $\{x; x > 0\}$.

In this case our random variable (or vector) is a *continuous random variable*.

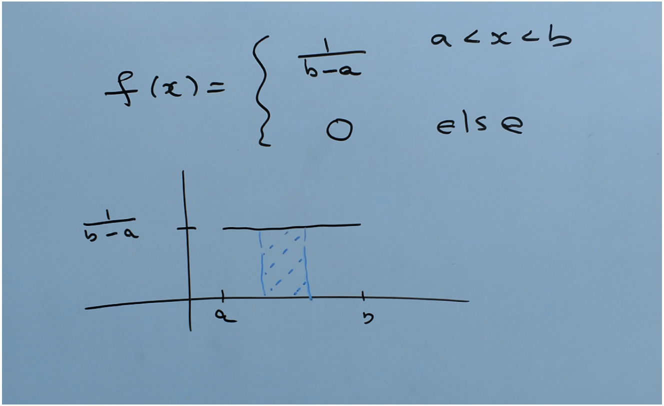
For continuous random variable we don't talk about the probability of a particular value, we can only talk about about the probability of a set.

We use the *probability density function* (*pdf*) f_x to specify the probability of a set A by

$$p(X \in A) = \int_A f_x(x)\, dx$$

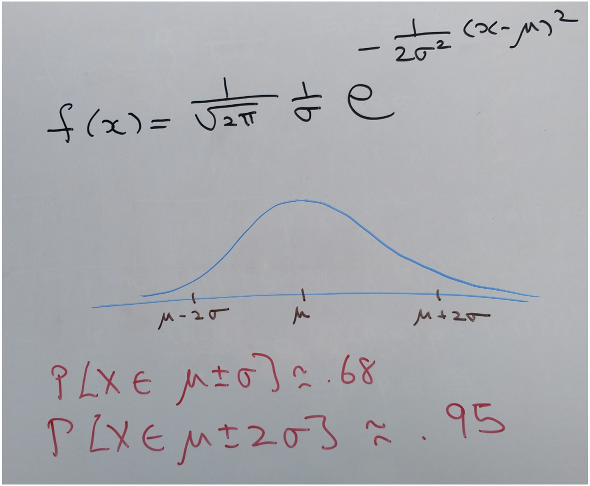
Example, the Uniform

The probability of any interval, is the area under the pdf over that interval !!!



We write $X \sim U(a, b)$.

Example, the Normal

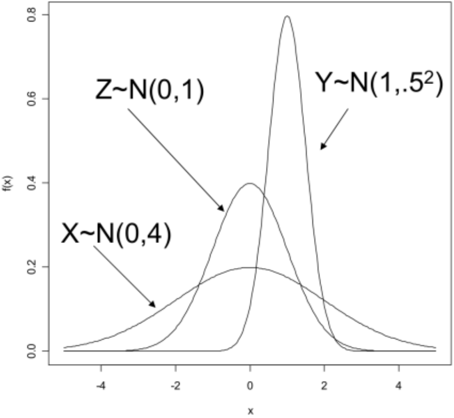


We write $X \sim N(\mu, \sigma^2)$.

The normal family has two parameters

μ: where the curve is centered

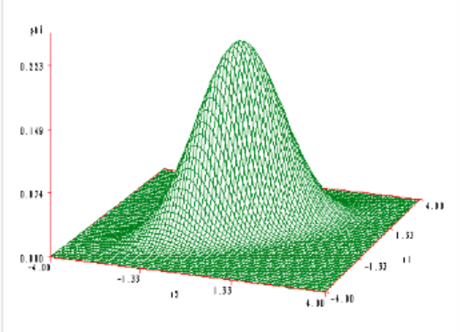
σ: how spread out the curve is



A small σ means the distribution is "tight" around μ !!

For more than one random variable we have the joint density:

$$P(Y_1 \in [a_1, b_1], Y_2 \in [a_2, b_2]) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(y_1, y_2)\, dy_1 dy_2$$



Basic Properties

The basic properities we had in the discrete case extend to the continous case:

$$f(y_1, y_2, y_3, \ldots, y_n) =$$
$$f(y_1)\, f(y_2 \mid y_1)\, f(y_3 \mid y_1, y_2) \ldots f(y_n \mid y_1, y_2, \ldots, y_{n-1})$$

If the Y_i are independent then

$$f(y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} f(y_i)$$

Margining out:

$$f(y_1) = \int f(y_1, y_2)\, dy_2$$

Conditional:

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f(y_2)}$$

Bayes theorem:

$$\begin{aligned} f(y_2 \mid y_1) &\propto f(y_2)\, f(y_1 \mid y_2) \\ &= f(y_1, y_2) \end{aligned}$$

2. Expection, Mean, Variance, Covariance

Let Y be a random variable (or vector).

Sometimes we want to summarize the possible values of some function of Y.

We use a probability weighted average:

Discrete:

$$E(h(Y)) = \sum h(y)p(y)$$

Continuous:

$$E(h(Y)) = \int h(y)\, f(y)\, dy$$

The key examples are the mean and variance of a univariate random variable.

The Mean:

$h(y) = y$

$$
\begin{aligned}
E(Y) &= \sum y\, p(y) \quad \text{(discrete)} \\
&= \int y\, f(y)\, dy \quad \text{(continuous)}
\end{aligned}
$$

We often write μ or μ_y for $E(Y)$.

The Variance:

$h(y) = (y - \mu)^2$.

$$\begin{aligned} Var(Y) &= \sum (y-\mu)^2 \, p(y) \quad \text{(discrete)} \\ &= \int (y-\mu)^2 \, f(y) \, dy \quad \text{(continuous)} \end{aligned}$$

We often write σ^2 or σ_y^2 for $Var(Y)$.

The Standard Deviation:

$$\sigma = \sqrt{(}\sigma^2)$$

is the *standard deviation*.

Note that σ has the same units as Y.

The variance and standard deviation summarize how close a random variable tends to be to its mean.

Example, the Bernoulli:

$X \sim \text{Bernoulli}(p)$ means:

x	$P(X = x)$
0	1-p
1	p

$E(X) = (1 - p) \times 0 + p \times 1 = p.$

$$\begin{aligned} Var(X) &= \sum_{i=1}^{n} P(x_i) \times [x_i - E(X)]^2 \\ &= (1 - p) \times (0 - p)^2 + p \times (1 - p)^2 \\ &= p(1 - p) \times [p + (1 - p)] \\ Var(X) &= p(1 - p) \end{aligned}$$

Example, the Normal:

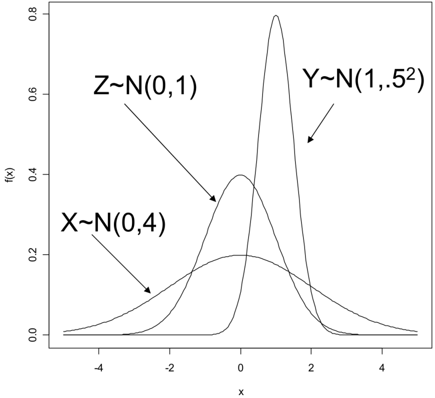
You can show that for $X \sim N(\mu, \sigma^2)$,

$$E(X) = \mu, \quad Var(X) = \sigma^2, \quad \sigma_X = \sigma.$$

The normal family has two parameters

μ: where the curve is centered

σ: how spread out the curve is



A small σ means the distribution is "tight" around μ !!

Covariance and Correlation:

The covariance and correlation are used to measure how much one random variable looks like a linear function of another.

Let $E(Y_1) = \mu_1$ and $E(Y_2) = \mu_2$.

$h(y_1, y_2) = (y_1 - \mu_1)(y_2 - \mu_2)$.

$$Cov(Y_1, Y_2) = E((Y_1 - \mu_1)(Y_2 - \mu_2)).$$

We might write $\sigma_{X,Y}$ for $Cov(X, Y)$, or σ_{12} for $Cov(Y_1, Y_2)$.

The Correlation

Let σ_i be the standard deviation of Y_i.

$$Cor(Y_1, Y_2) = \frac{\sigma_{12}}{(\sigma_1 \, \sigma_2)}.$$

The covariance divided by the product of the the standard deviations.

We might write ρ_{XY} for $Cor(X, Y)$ for ρ_{12} for $Cor(Y_1, Y_2)$.

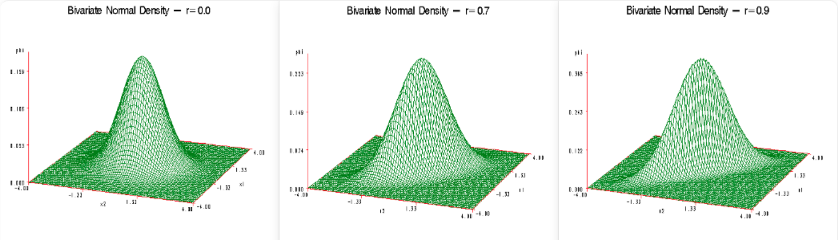
Key Property of Correlation:

$$-1 \leq \rho_{X,Y} \leq 1$$

The correlation is always between 1 and -1.

The closer the correlation is to 1, the more $Y \approx a + bX$ with $b > 0$.

The closer the correlation is to -1, the more $Y \approx a + bX$ with $b < 0$.


Bivariate Normal Density — r=0.0
Bivariate Normal Density — r=0.7
Bivariate Normal Density — r=0.9

Suppose X and Y are independent.

Then,

$$\begin{aligned} \sigma_{XY} &= E((X - \mu_X)(Y - \mu_y)) \\ &= E(X - \mu)E(Y - \mu) \\ &= 0 \times 0 = 0 \end{aligned}$$

Independent $\Rightarrow \rho_{XY} = 0$, but not the other way around.