Class Notes

Statistical Computing & Machine Learning

Class 12

Review

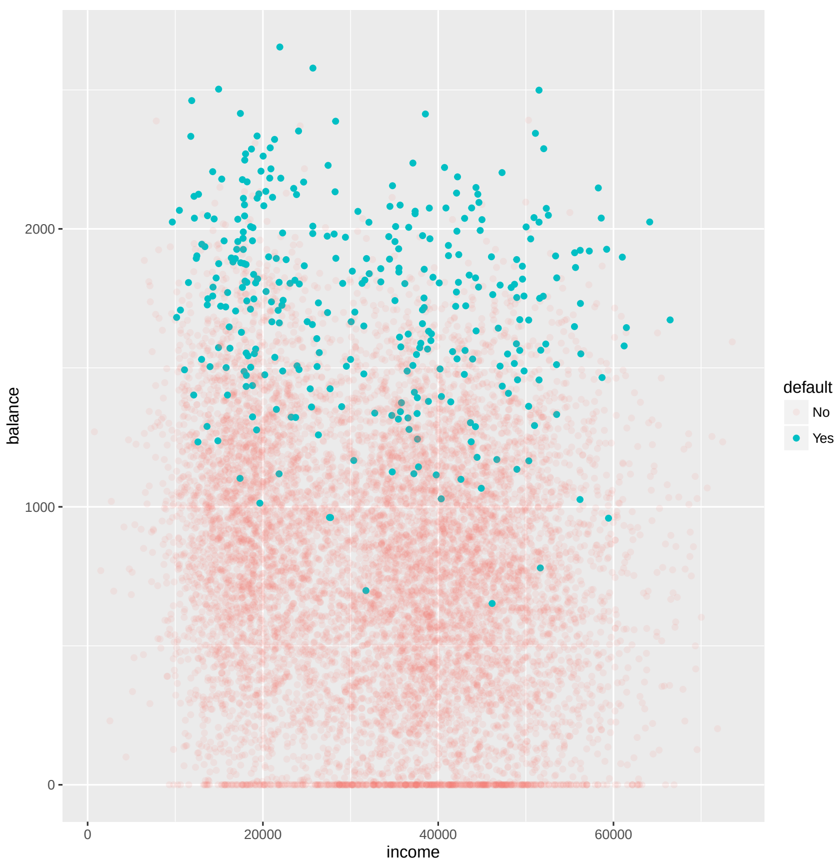
Situation: Build a classifier. We measure some features and want to say which group a case refers to.

Specific example: Based on the `ISLR::Default` data, find the probability of a person defaulting on a loan given their `income` and `balance`.

```
names(Default)
```

```
## [1] "default" "student" "balance" "income"
```

```
ggplot(Default,
       aes(x = income, y = balance, alpha = default, color = default)) +
  geom_point()
```



We were looking at the likelihood: prob(observation | class)

Note: Likelihood itself won't do a very good job, since defaults are relatively uncommon. That is, p(default) $\ll$ p(not).

A standard (but not necessarily good) model of a distribution is a multivariate Gaussian.

Univariate Gaussian

$$p(x) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}}_{\textit{Normalization}} \underbrace{\exp(-\frac{(x-m)^2}{2\sigma^2})}_{\textit{Shape}}$$

Imagine that we have another variable $z = x/3$. Geometrically, z is a stretched out version of x, stretched by a factor of 3. The distribution is

$$p(z) = \underbrace{\frac{1}{\sqrt{2\pi(3\sigma)^2}}}_{\textit{Normalization}} \underbrace{\exp(-\frac{(x-m)^2}{2(3\sigma)^2})}_{\textit{Shape}}$$

Note how the normalization changes. $p(z)$ is broader than $p(x)$, so it must also be shorter.

Uncorrelated bivariate gaussian

For independent RVs x and y, p(xy) = p(x)p(y). Show that the normalization is $\frac{1}{2\pi\sigma_x\sigma_y}$.

The sigmas multiply in the normalization, like the area of something being stretched out in two orthogonal directions.

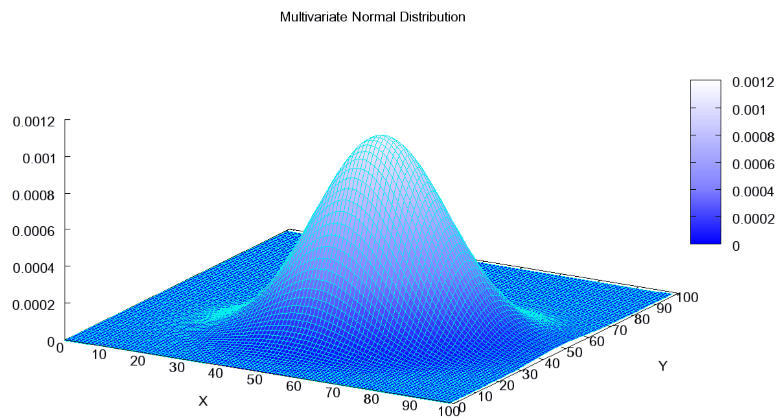
Bivariate normal distribution with correlations

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right)$$

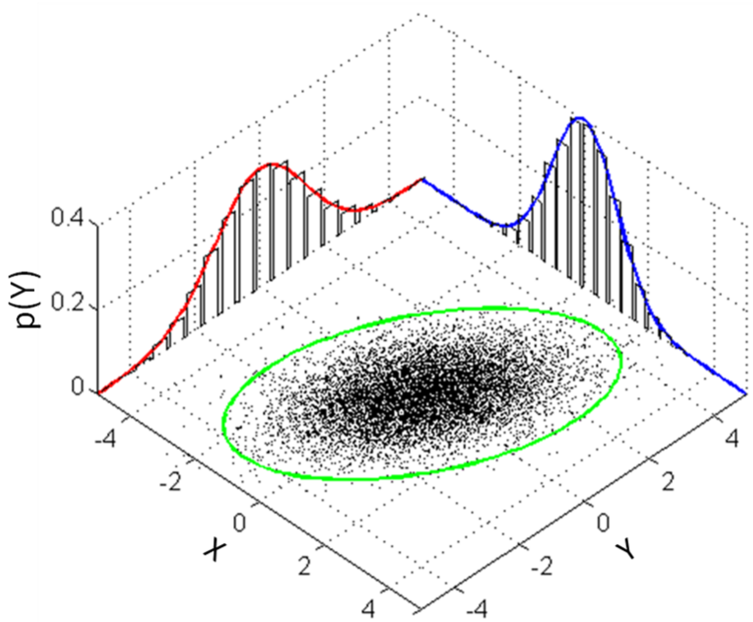
If $\rho > 0$ and x and y are both above their respective means, the correlation term makes the result *less* surprising: a larger probability.

Or, for multiple dimensions

$$(2\pi)^{-\frac{k}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$



As an amplitude plot



Showing marginals and 3-σ contour

Determinant gives area of a parallelogram

Look at the stretching that goes on

Proof without words:
A 2×2 determinant is the area of a parallelogram