1 Computations with random variables

It is not rare, that random variables are transformed through simple arithmetic operations. Consider ie. instead of being interested in the random variable X modelling the sum of two independent rolls with a die, we could be looking at the random variable $Y = 2X + 1$, which for each the values that can be obtained are determined by the sum of the two die rolls and are transformed in a way, that they are multiplied by 2 and added 1 to. Instead of reformulating the entire range Y is defined on and computing the probability distribution through the PMF and CDF, we would like to be able to compute the expectation of this new random variable fast and conveniently.

Luckily, the single-point summaries of distributions, the expectation and variance of random variables have nice properties when it comes to transformations.

1.1 Transforming (scaling and shifting) random variables

How does the mean and the standard deviation get affected if we were to add to the random variable, or if we were to scale the random variable?

X is our random variable. If a constant were added:

$Y = X + k$

Then the mean will the affected by a addition of k:

$\mu_Y = \mu_X + k$

BUT the standard deviation of Y will not be affected:

$\sigma_Y = \sigma_X$

If X is scaled by some constant k:

$Z = k \cdot X$

The both the mean and the standard deviation will be scaled by k:

$\mu_Z = k \cdot \mu_X$ and $\sigma_Z = k \cdot \sigma_X$

1.1.1 Change-of-Variable Formula

For any discrete random variable X and some defined function we can obtain the expected value of X through the following formula:

$$E[g(X)] = \sum_{x_k \in R_X} g(x_i)P(X = x_k)$$

And for the continuous case, we similarly observe:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

The change-of-variable formula is the most generic way of recomputing the expectation for a transformed random variable, since it works with any kind of transformation function g. This is sensible, since all it does is applying the function to the former range of the random variable and incorporating this new range into the regular formulas for computing expectations.

1.1.2 Linearity of Expectation (LOTUS)

However, for linear transformations of the type $f(x) = ax + b$, we can use the LOTUS-formula to show that:

$$E[ax + b] = a \cdot E[X] + b$$

This allows us to recompute the expectation of a linearly transformed random variable simply by transforming the computed value of the expectation itself. Note that since $E[aX] = E[X_1] + E[X_2] + \ldots + E[X_a]$, then the total expectation of independent trials of the same random phenomena modelled through X is simply the sum of each single expectation. Adding a constant to some random variable $E[X + b]$ simply shifts the range R_X by the constant, but leaves the probability distribution untouched. Then, the expectation also simply shifts by that constant $E[X] + b$.

1.1.3 Jensen Inequality

The inequality says, that if you are given a convex function and a random variable X then this inequality holds:

$$f(E[X]) \leq E[f(x)]$$

the LHS of the formula is telling you how to calculate this value. You draw n values of X and you calculate their average (of n), then you take that average and you pass it into our convex function, which is given us back a number. That number is the LHS (output of the average input). On the RHS, you draw n samples of X, but this time we pass each into the function, giving us many outputs, and then we take the average of those outputs, giving us the RHS (the average output).

So Jensen Inequality states, that the output of the average input is less than (or equal to) the average output.

A convex function

When a function is a convex function, then the epigraph (the space of points above the funciton) is a convex set. Which means if you pick any two points in that epigraph and then draw a line between them, then every point along that line is also within that epigraph.
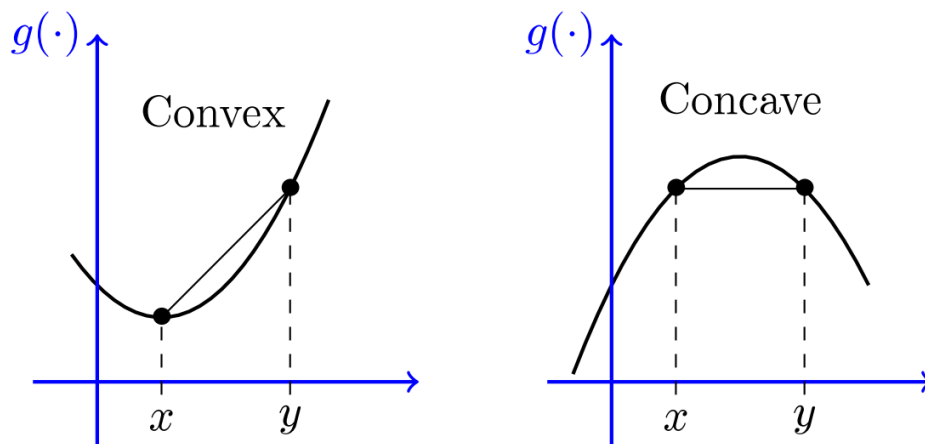

Figure 1: Concave and convex functions