Sampling from a Probability Distribution

CSCI 1440/2440

2025-10-08

We introduce quantiles, and show that sampling a quantile uniformly at random, and then inverting a distribution F at that sample value, is equivalent to sampling from F itself.

1 *Quantiles*

A **quantile** $q \in [0, 1]$ is the relative strength of a value $v \in T$:

$$q(v) = 1 - F(v).$$

Likewise,

$$F(v) = 1 - q(v).$$

The value $F(v)$ is the probability that the value of a random draw from the distribution F is less than or equal to v. Accordingly, the quantile v is the probability that the value of a random draw is greater than v. Thus, lower quantiles correspond to higher types/values, and higher quantiles correspond to lower types/values.

To map a quantile back to a value, we invert F at $1 - q$:

$$v(q) = F^{-1}(1 - q).$$

Perhaps the most famous quantile is the **median**, which is the value of a CDF F inverted at 1/2. The 10th% quantile corresponds to the value at the 90th% percentile, meaning the value below which 90% percent of the data lie. (And likewise, for the kth quantile and the $100 - k$th percentile.)

N.B. Quantiles are more often defined as $F(v)$, rather than $1 - F(v)$. Our choice is natural in the context of this lecture, whose subject is posted-price revenue maximization, because the probability of a sale is the probability that a buyer's value (i.e., a draw from F) exceeds the posted price.

2 *Probability Integral Transform*

Let X be a random variable with arbitrary CDF F. Now define the random variable $Y = F(X)$. Note that the range of Y is $[0, 1]$. One interesting question is, how is Y distributed? That is, *what is the CDF of an arbitrary CDF?* The answer to this question is straightforward to derive, but surprising nonetheless.

For $0 \leq y \leq 1$,

$$\begin{aligned}\Pr(Y \leq y) &= \Pr(F(X) \leq y) \\ &= \Pr(X \leq F^{-1}(y)) \\ &= F(F^{-1}(y)) \\ &= y.\end{aligned}$$

Since $\Pr(Y \leq y) = y$, $Y = F(X)$ is uniformly distributed on $[0, 1]$! That is, *the CDF of an arbitrary CDF is uniform*. Finally, since a quantile associated with a value is 1 less the CDF at that value, it follows that *quantiles are likewise uniformly distributed*.

3 *An Aside: Inverse Transform Sampling*

Let's apply this same logic to the random variable X, defined as the inverse of the uniform distribution for some cdf F (i.e., $X = F^{-1}(U)$):

$$\begin{aligned}Pr(X \leq x) &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(U \leq F(x)) \\ &= F(x).\end{aligned}$$

So $X = F^{-1}(U)$ is distributed according to F.

An important practical consequence of this observation is a process for sampling from an arbitrary CDF:

1. Sample from the uniform distribution to obtain a value of $F(x)$.
2. Apply F^{-1} to that sample to recover $x = F^{-1}(F(x))$, which is necessarily distributed according to F.

This process is called **inverse transform sampling**.

Example 3.1. The uniform distribution on $[0, k]$ has CDF $F(x) = x/k$ and inverse $F^{-1}(y) = ky$. Using inverse transform sampling, we can sample y from $U[0, 1]$, and then obtain as our sample $F^{-1}(y) = F^{-1}(F(x)) = F^{-1}(x/k) = k(x/k) = x$. From the transform, we know that this is equivalent to sampling x from $U[0, k]$.

Example 3.2. The exponential distribution has CDF $F(x) = 1 - e^{-\lambda x}$. We invert this CDF as follows:

$$\begin{aligned}F(x) &= 1 - e^{-\lambda x} \\ e^{-\lambda x} &= 1 - F(x) \\ -\lambda x &= \ln(1 - F(x)) \\ x &= -\ln(1-F(x))/\lambda \\ F^{-1}(y) &= -\ln(1-y)/\lambda.\end{aligned}$$

Now, using inverse tranform sampling, we can sample from the exponential distribution by first sampling a value y from $U[0,1]$, and then obtain as our sample $F^{-1}(y) = F^{-1}(F(x)) = F^{-1}(1 - e^{-\lambda x}) = {}^{-\ln\left(1-(1-e^{-\lambda x})\right)}/_{\lambda} = {}^{\lambda x}/_{\lambda} = x$. From the transform, we know that this is equivalent to sampling x from the exponential distribution.