Stat 5101 Lecture Slides
Deck 4

Charles J. Geyer

School of Statistics

University of Minnesota

Quantiles

If X is a random variable and $0 < q < 1$, then the *q-th quantile* of X (or of the distribution of X) is any number x such that

$$\Pr(X \leq x) \geq q \quad \text{and} \quad \Pr(X \geq x) \geq 1 - q$$

If X is a continuous random variable having distribution function F, then this simplifies to

$$F(x) = q.$$

If X is a discrete random variable having distribution function F, then this simplifies to

$$F(y) \leq q \leq F(x), \qquad y < x.$$

The distribution function of the $\text{Bin}(3, 1/2)$ distribution is

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/8, & 0 \le x < 1 \\ 1/2, & 1 \le x < 2 \\ 7/8, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

and

$\quad$ 0 is the unique q-th quantile for $0 < q < 1/8$

$\quad$ x is a 1/8-th quantile for $0 \le x \le 1$

$\quad$ 1 is the unique q-th quantile for $1/8 < q < 1/2$

$\quad$ x is a 1/2-th quantile for $1 \le x \le 2$

and so forth.

Quantiles (cont.)

For a continuous random variable whose support is an interval, quantiles are uniquely defined.

Suppose X is an $\text{Exp}(\lambda)$ random variable, then the q-th quantile is the solution for x of

$$F(x) = 1 - e^{-\lambda x} = q$$

which has the solution

$$x = -\frac{1}{\lambda} \cdot \log(1 - q)$$

Quantiles (cont.)

Certain special quantiles have other names.

The 1/2-th quantile is also called the *median*.

The 1/4-th and 3/4-th quantile are also called the *lower quartile* and *upper quartile*, respectively.

If $100q$ is an integer, the q-th quartile is also called the $100q$-th percentile. (This terminology is frowned upon in serious probability theory, just like all use of percentages.)

Quantile Functions

Although quantiles are not necessarily unique, it is mathematically convenient to define a function that gives for each q a q-th quantile.

If F is a distribution function, the corresponding *quantile function* G is

$$G(q) = \inf\{\, x \in \mathbb{R} : F(x) \geq q \,\}, \qquad 0 < q < 1$$

If the q-th quantile is unique, then it is $G(q)$.

If the q-th quantile in non-unique, then $G(q)$ is the smallest q-th quantile.

Quantile Functions (cont.)

For the Exp(λ) distribution, the DF is

$$F(x) = (1 - e^{-\lambda x})I_{[0,\infty)}(x), \qquad -\infty < x < \infty$$

and the quantile function is

$$G(q) = -\frac{1}{\lambda} \cdot \log(1 - q), \qquad 0 < q < 1$$

For the Bin$(3, 1/2)$ distribution, the DF is

$$
F(x) = \begin{cases}
0, & x < 0 \\
1/8, & 0 \le x < 1 \\
1/2, & 1 \le x < 2 \\
7/8, & 2 \le x < 3 \\
1, & 3 \le x
\end{cases}
$$

and the quantile function is

$$
G(q) = \begin{cases}
0, & 0 < q \le 1/8 \\
1, & 1/8 < q \le 1/2 \\
2, & 1/2 < q \le 7/8 \\
3, & 7/8 < q < 1
\end{cases}
$$

Quantile Functions (cont.)

If the support of a random variable is the whole real line, then the DF is invertible and the quantile function is its inverse.

An example is the standard normal distribution, whose DF is conventionally denoted Φ and whose quantile function is conventionally denoted Φ^{-1}. Neither has an expression in terms of other known functions. If you want to use either, you must use a computer.

Quantile Functions (cont.)

If the support of a random variable is an interval, then the restriction of the DF to the interior of this interval is invertible and the quantile function is its inverse.

Examples are all the other brand name continuous distributions. For many of these, such as the gamma and beta distributions, neither the DF nor its inverse has an expression in terms of other known functions. If you want to use either, you must use a computer.

Best Prediction

A long time ago we learned that $\mu = E(X)$ is the best prediction of the value of X if "best" means minimizing expected squared error, that is, μ minimizes

$$a \mapsto E\{(X - a)^2\}.$$

But what if we don't accept the minimizing expected squared error criterion?

Also, what if the mean doesn't exist? Then what is the best prediction?

Best Prediction (cont.)

If X is a random variable, and a is a prediction of the value of X before it is observed, then

$$E\{|X - a|\},$$

the *absolute error* of the prediction, is minimized when a is any median of the distribution of X.

Best Prediction (cont.)

Suppose m is a median of X. If $x < m \leq a$, then

$$|x - a| - |x - m| = (a - x) - (m - x) = a - m.$$

If $m \leq x \leq a$, then

$$|x - a| - |x - m| = (a - x) - (x - m) = a + m - 2x$$

If $m \leq a < x$, then

$$|x - a| - |x - m| = (x - a) - (x - m) = -(a - m)$$

Best Prediction (cont.)

Summarizing, in case $m \leq a$

$$|x - a| - |x - m| = \begin{cases} a - m, & x < m \\ a + m - 2x, & m \leq x \leq a \\ -(a - m), & x > a \end{cases}$$

and

$$|x - a| - |x - m| \geq -(a - m) + 2(a - m)I_{(-\infty, m]}(x)$$

Hence

$$E\{|X - a|\} - E\{|X - m|\} \geq -(a - m) + 2(a - m)\Pr(X \leq m) \geq 0$$

Best Prediction (cont.)

A similar analysis of the case $a \leq m$ shows that

$$E\{|X - a|\} - E\{|X - m|\} \geq 0$$

whatever a may be.

Best Prediction (cont.)

In summary, any median minimizes the function

$$a \mapsto E\{|X - a|\}$$

which is called expected absolute error. Hence the median is the "best" prediction in this sense.

If first moments don't exist, then we can say the median minimizes

$$a \mapsto E\{|X - a| - |X - m|\}$$

which we have seen is a bounded function that always has expectation.

Best Prediction (cont.)

The mean is the best prediction if minimizing expected squared error is the criterion.

The median is the best prediction if minimizing expected absolute error is the criterion.

Something else would be the best prediction if some other criterion were used. But we leave this theme here.