16. Quantile-Quantile (QQ) Plots II
The Four Patterns of Quantiles-Quantile (QQ) Plots

8:23 / 8:23 1.0x

Again, this won't change, because I
can go from one to the other by just
changing the standard deviation and the mean.
And therefore all I'm going to have to do
is to just change the line.
So if you see something that looked like an S shaped,
it's going to be a uniform distribution.
You just don't know which one.

End of transcript. Skip to the start.

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

QQ Plots and Tails of Distributions

1/2 points (graded)
Suppose we have an ordered data set $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. Let Φ denote the cdf of the distribution $\mathcal{N}(0,1)$. Recall that the QQ plot of this data against the normal distribution is given by plotting the ordered pairs

$$\left(\Phi^{-1}(1/n), X_{(1)}\right), \left(\Phi^{-1}(2/n), X_{(2)}\right), \ldots, \left(\Phi^{-1}(i/n), X_{(i)}\right), \ldots, \left(\Phi^{-1}((n-1)/n), X_{(n-1)}\right).$$

The x-values above

$$\Phi^{-1}(1/n), \Phi^{-1}(2/n), \ldots, \Phi^{-1}(i/n), \ldots, \Phi^{-1}((n-1)/n)$$

are referred to as the **theoretical quantiles** , and the y-values above given by the reordered sample

$$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$$

are referred to as the **empirical quantiles** .

We say that a distribution $\mathbf{P}$ has a **heavier right tail** if

$$P(X \geq t) \geq P(Y \geq t) \quad \text{for } t > 0 \text{ sufficiently large,}$$

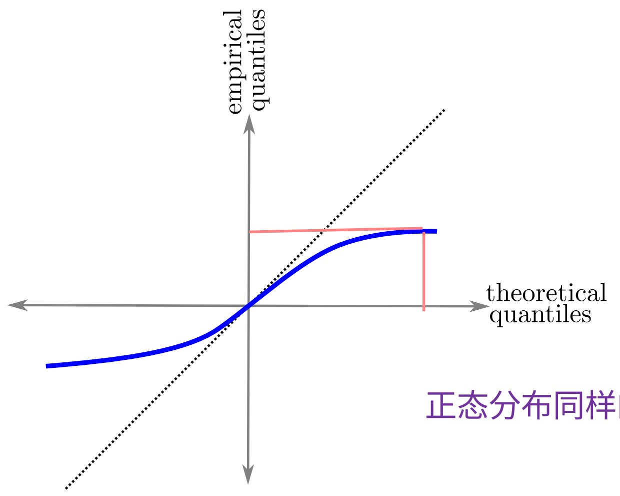
where $X \sim P$ and $Y \sim Q$. (Otherwise, $\mathbf{P}$ is said to have a **lighter** right tail than Q.)

Similarly, we say that $\mathbf{P}$ has a **heavier left tail** if

$$P(X \leq -t) \geq P(Y \leq -t) \quad \text{for } t > 0 \text{ sufficiently large,}$$

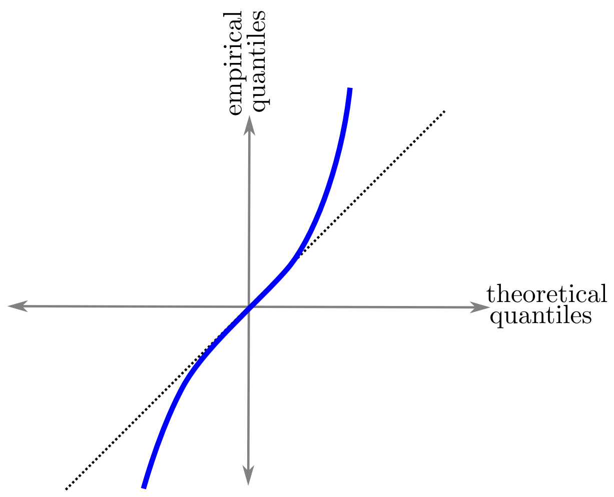
where $X \sim P$ and $Y \sim Q$. (Otherwise, $\mathbf{P}$ is said to have a **lighter** left tail than Q.)

Which of the following QQ plots has a lighter right tail than that of a standard Gaussian?

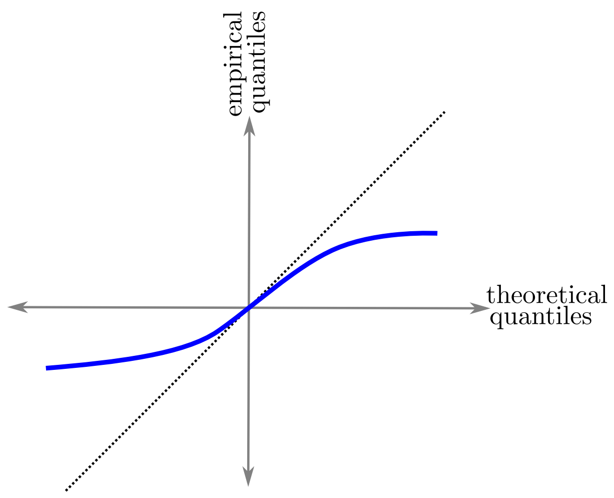


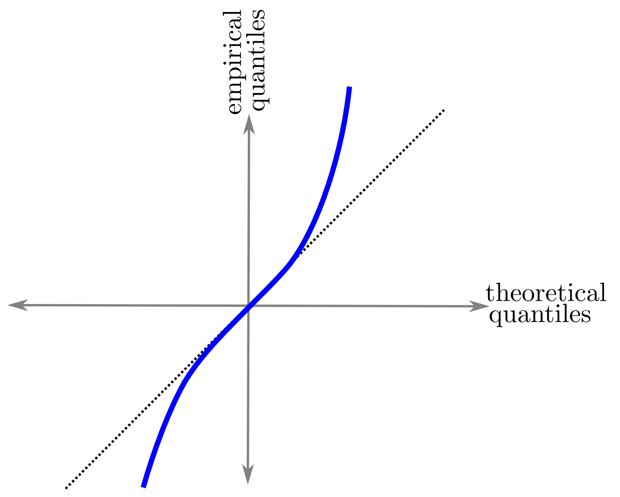
重尾更多的质量在尾部，同样的概率，quantile也就是Tn就会更大
轻尾更多的质量在中间，同样的概率，quantile也就是Tn就会更小

正态分布同样的概率，quantile也就是Tn更大，所以Y更轻



Which QQ plot has a lighter left tail than that of a standard Gaussian?



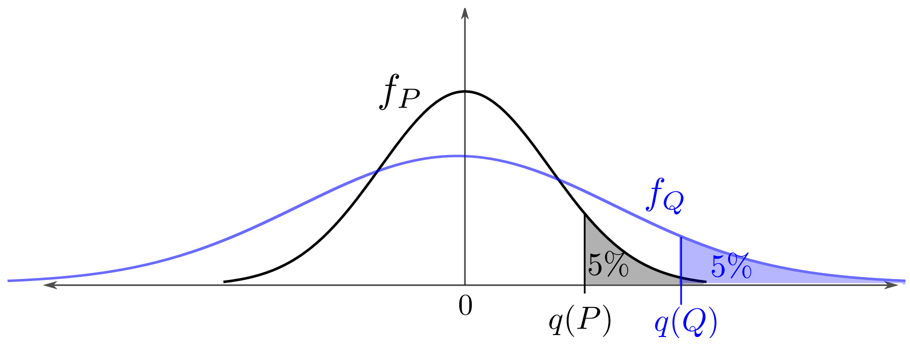


Solution:

Question 1: The correct answer to the first question is the first plot. According to the first QQ plot, we can see that the empirical quantiles on the right are smaller than the theoretical (*i.e.*, quantiles of a standard Gaussian) on the right. To understand this, we consider a thought experiment comparing two continuous distributions $\mathbf{P}$ and Q.

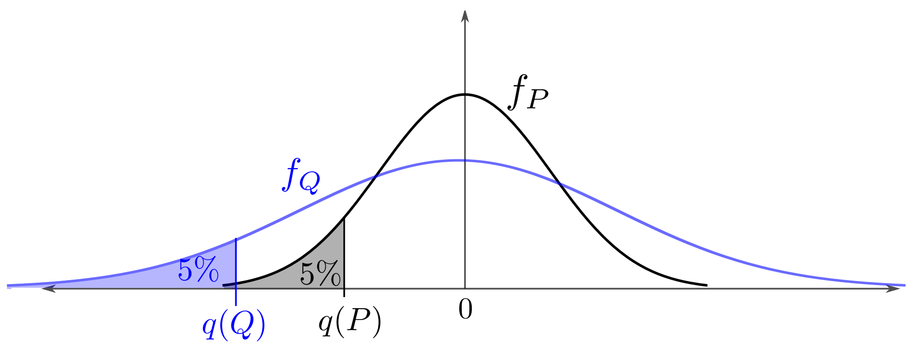
Remark: The empirical distribution is of course a discrete distribution, but you should convince yourself that the the logic below still holds in that case as well.

Suppose that the η quantiles of $\mathbf{P}$, denoted by q_η, are much smaller than the η quantiles of Q, denoted by q'_η. This is illustrated in the figure below where we set $\eta = 5\%$.



We can go far out into the tails of Q and still have area under the curve to the right to be 0.05. However, from the graphic, if we look at the area to the right of $q_{0.05}$ (the 5% quantile for Q) under the density of $\mathbf{P}$, this will be significantly smaller than 0.05. Hence, the right tails of $\mathbf{P}$ are **lighter** than the right tails of Q.

Question 2: The correct answer to the second question is the first plot. We see that the empirical quantiles on the left are larger than the theoretical quantiles on the left. You should convince yourself that this is the 'mirror image' of the thought experiment above with $\mathbf{P}$ and Q. In this case as well, $\mathbf{P}$ has lighter left tails than Q.



Submit You have used 1 of 1 attempt

Answers are displayed within the problem

QQ Plot for a Student's T Distribution

0/1 point (graded)

Refer to the images in the slide "Quantile-Quantile (QQ) Plots (3)" which show data drawn from a student's T distribution (for $n = 10, 500, 100, 1000, 5000, 10000$) compared to the standard normal $\mathcal{N}(0, 1)$.

Upon examining the QQ plots in that slides, would you conclude that the tails (both left and right) of t_n are **heavier** or **lighter** than those of the standard normal distribution?

- Lighter
- Heavier ✔

Solution:

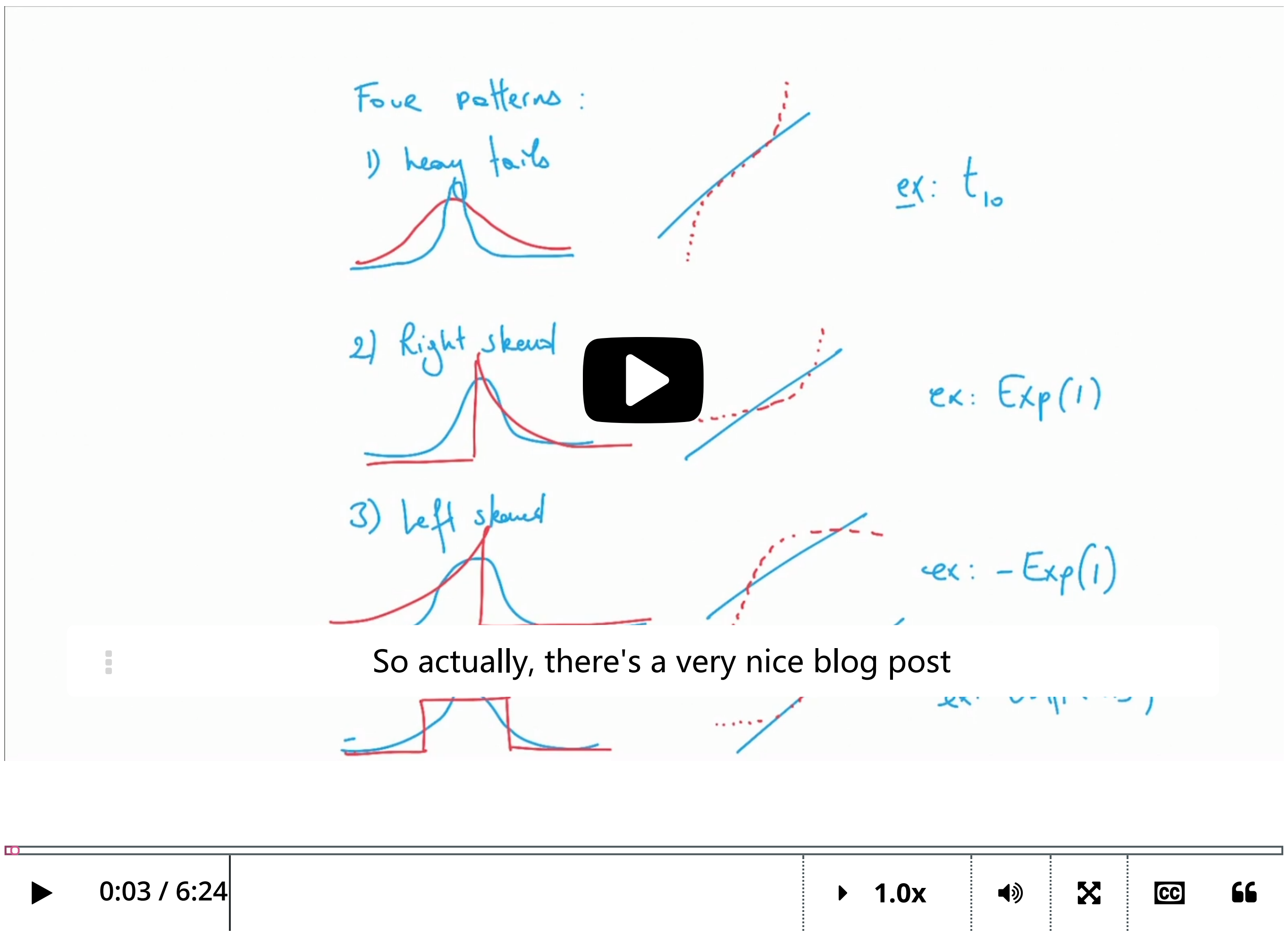
The correct response is "Heavier." We see that for all values of n shown in this slide, the QQ plot lies **below** the line $y = x$ to the left and **above** the line $y = x$ to the right.

Recall that the left tails are heavier than those of $\mathcal{N}(0, 1)$ if the QQ plot lies **below** the line $y = x$ (see the solution to the previous problem for an explanation). Moreover, the right tails are heavier than those of $\mathcal{N}(0, 1)$ if the QQ plot lies **above** the line $y = x$ (see the solution to the previous problemn).

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

A Blog Post on Quantiles-Quantile (QQ) Plots



The line is missing, so you just throw QQ line.

And you might decide that you want

it to be blue and have thickness too,

but you don't really have to do this.

And so you're going to get the QQ

plot that I showed you before.

And on top of it, it's going to super imprint this line.

OK?

And here, as you can see, so here, it's testing a Gaussian.

So let's not really pay attention

to too much about the slope.

And the intercept, this should actually

be very close to 1 and 0, respectively.

But let's see for some other example what's happening.

OK?

So now, I move on to something which

is skewed to the right or skewed right or right skewed,

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Discussion

Show Discussion

Topic: Unit 4 Hypothesis testing:Lecture 16: Goodness of Fit Tests Continued: Kolmogorov-Smirnov test, Kolmogorov-Lilliefors test, Quantile-Quantile Plots / 16. Quantile-Quantile (QQ) Plots II