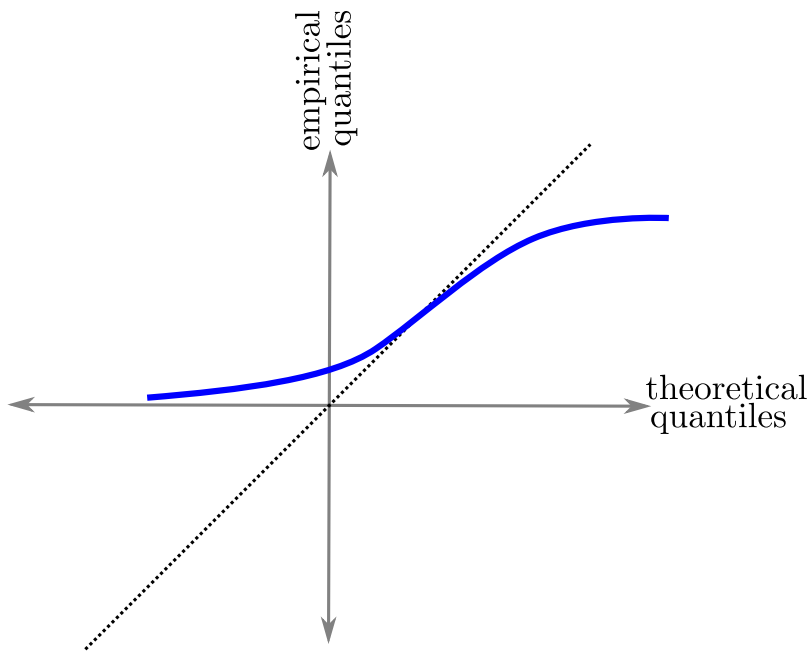
17. Quantile-Quantile (QQ) Plots II

Matching a Distribution to a QQ Plot

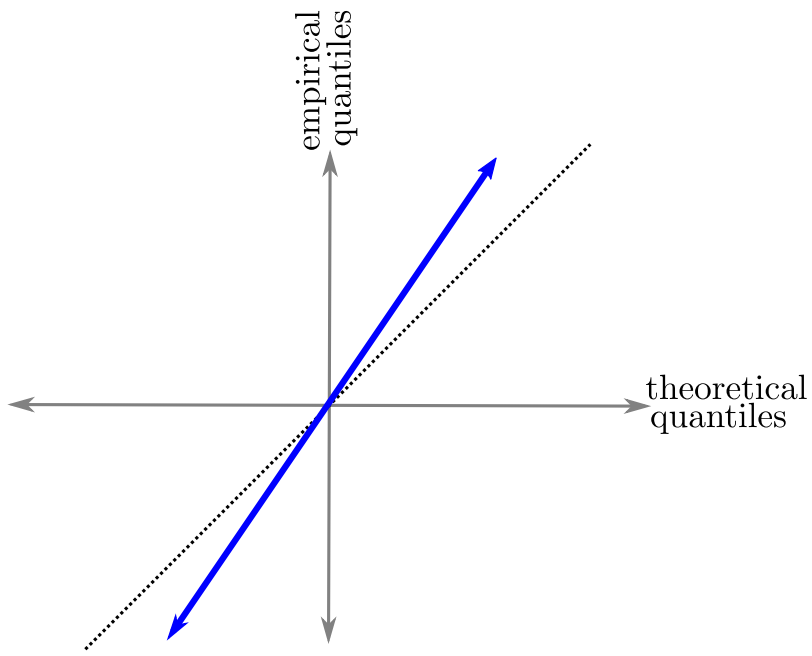
4/4 points (graded)

Consider an iid sample $X_1, X_2, \ldots, X_n \overset{iid}{\sim} P$ that has been reordered as $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. In each image below, we have chosen a different distribution for $\mathbf{P}$ and compared the empirical quantiles to the standard Gaussian quantiles using a QQ plot. Assume that n is large enough so that the QQ plot starts to look like a continuous curve.

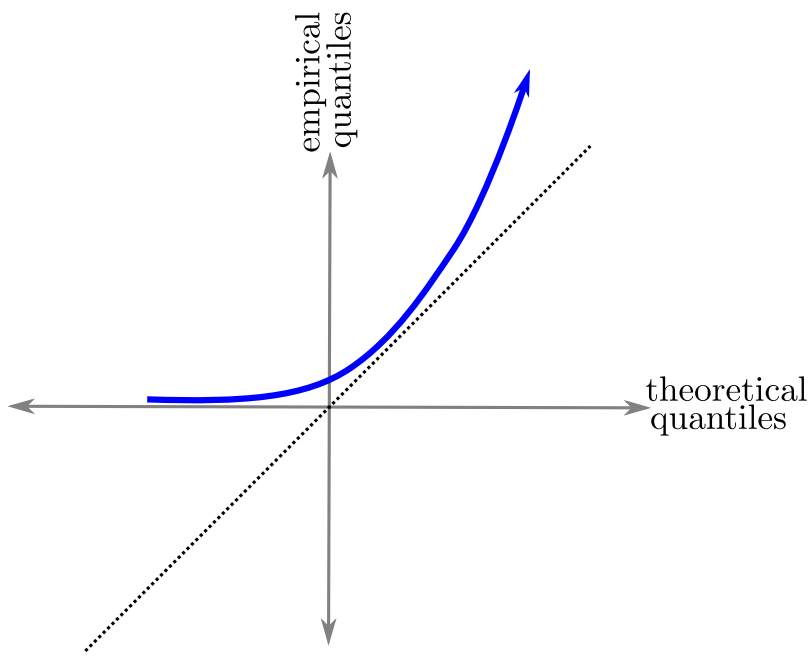
For each plot, match the QQ plot with the correct distribution for $\mathbf{P}$. Each possible distribution will be an answer choice exactly once, so you should use the process of elimination.



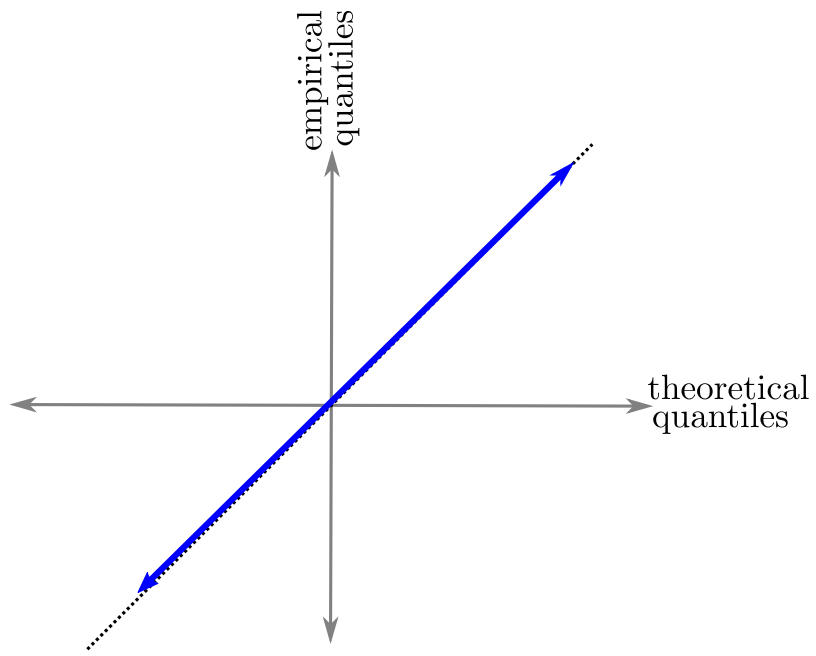
- Uniform on $[0, 1]$ ✔
- Exponential with mean 1: $\mathbf{Exp}\,(1)$
- Standard Gaussian $\mathcal{N}\,(0, 1)$
- Gaussian with variance 10: $\mathcal{N}\,(0, 10)$



- Uniform on $[0, 1]$
- Exponential with mean 1: $\mathbf{Exp}\,(1)$
- Standard Gaussian $\mathcal{N}\,(0, 1)$
- Gaussian with variance 10: $\mathcal{N}\,(0, 10)$ ✔



- Uniform on $[0, 1]$
- Exponential with mean 1: $\mathbf{Exp}\,(1)$ ✔
- Standard Gaussian $\mathcal{N}\,(0, 1)$
- Gaussian with variance 10: $\mathcal{N}\,(0, 10)$



- Uniform on $[0, 1]$
- Exponential with mean 1: $\mathbf{Exp}\,(1)$
- Standard Gaussian $\mathcal{N}\,(0, 1)$ ✔
- Gaussian with variance 10: $\mathcal{N}\,(0, 10)$

Solution:

Question 1: The distribution for this QQ plot is **uniform on $[0, 1]$**. Since the support for this distribution is $[0, 1]$, the empirical quantiles $X_{(1)}, X_{(2)}, \ldots, X_{(n)} \in [0, 1]$. Since there is nothing plotted outside of the interval $[0, 1]$ on the y-axis, we see that the support is restricted to this interval. This implies by process of elimination that the sample was generated from a uniform distribution.

Question 2: The dist ribution for this QQ plot is **Gaussian with variance 10: $\mathbf{N}\,(0, 10)$**. The QQ plot is a straight line, which suggests that the data is drawn from a Gaussian distribution. However, the slope is significantly larger than that of the the line $y = x$. Hence, the tails of $\mathbf{P}$ must be **heavier** than those of $\mathcal{N}\,(0, 1)$. A larger variance results in heavier tails, so by process of elimination, the data must be generated from $\mathcal{N}\,(0, 10)$.

Question 3: The distribution for this QQ plot is **Exponential with mean 1 : $\mathbf{Exp}\,(1)$**. Note that the exponential distribution is supported on $[0, \infty)$. Hence, the QQ plot will not go below the line $y = 0$. Moreover, the exponential distribution has **heavier** tails than those of $\mathcal{N}\,(0, 1)$, so we expect the QQ plot to be above the line $y = x$, which is indeed the case here.

Question 4: The distribution for this QQ plot is **Standard Gaussian $\mathbf{N}\,(0, 1)$**. Observe that the QQ plot lies very close to the line $y = x$, so this suggests that the the data is distributed as $\mathcal{N}\,(0, 1)$. By process of elimination, we conclude that the data must have been generated from a standard Gaussian.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 4 Hypothesis testing:Lecture 16: Goodness of Fit Tests Continued: Kolmogorov-Smirnov test, Kolmogorov-Lilliefors test, Quantile-Quantile Plots / 17. Quantile-Quantile (QQ) Plots II