Questions
1. When did Pascal and Laplace write their great works on probability? How about Bayes?
   Pascal: mid 17th century (1654)
   Laplace: early 19th century (1812)
   Bayes: mid 18th century (1736)

2. What formula relates the probability of two events to the probability of each event, if they are independent?
   P(A and B) = P(A) * P (B)

3. What defines independence?
   One occurrence cannot affect the probability of another occurrence.

4. In your own words, what is the frequentist interpretation of probability?
   The expected percentage of an event occurring (e.g., you would expect a die to land on three 16.67% of the time) if the event were to occur an infinite amount of times.

5. In your own words, what is the Bayesian interpretation of probability?
   Making the most reasonable prediction based on given previous information (e.g., Team A has beaten Team B in 75% of their games, so one would expect Team A to win the next game).

6. In Bayesian statistics, what are prior and evidence and posterior? Write Bayes' theorem and define the symbols you use, perhaps with an example. How does it relate to the scientific method of a cycle from hypothesis to experiment and on to better hypothesis?
   Prior: a probability assigned to an event using any knowledge of the event and its probability before any evidence is acquired.
   Evidence: information gathered through testing, typically empirically.
   Posterior: a probability assigned to an event taking into account both priors and evidence gathered after the prior was assigned.
   P(A|B) = P(B|A)*P(A)/P(B) (P = probability, A = one event, B = different event, | conditional probability)
   You make an initial hypothesis based on assumptions or previous knowledge; you then gather evidence and more information, and then adjust your hypotheses for further testing as needed

7. What is a probability distribution? a probability density? (hint: discrete vs. continuous). What are the units of p(T), the probability density of temperature of some object at some time?
   Discrete: A histogram that represents the probability of an occurrence of an event with multiple outcomes.

Continuous: A function that represents the probability of an occurrence of an event with multiple outcomes.
Units: the inverse of whatever time is being used

8. Using the term likelihood for probability density, what is the maximum likelihood value of T if p(T) is not symmetric -- the mean, the median, or the mode? Sketch a nonsymmetric distribution and indicate these 3 different measures of its central tendency.
   The mode
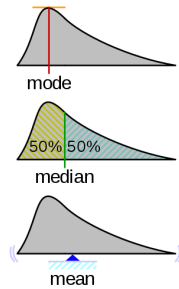
9. Consider a uniform distribution over [1,2]. What is its first moment? What are its first four central moments?
   First: 1.5
   First Central: 0
   Second Central: 1/12
   Third Central: 0
   Fourth Central: -1.2

10. What fundamental mathematical operation (addition, subtraction, multiplication, division) creates the Normal distribution according to the Central Limit Theorem? Can you think of reasons it is so commonly observed? That is, can you name some natural processes that mimic that mathematical operation?
    Sum; Many things in nature are the "sum" of many different independent factors, such that they can be thought of as truly random and therefore follow a normal distribution (but not the same one)
    Many human features (e.g., height) follow normal distributions

11. In your own words, what is a test statistic?
    A contrived value that can be used to verify or reject a hypothesis with a single value

12. What is the Z-test (based on the Z-statistic)? What are the one-tailed (or one-sided) p-values for Z values of 1, 2, 3? (sometimes called one-sigma, two-sigma, three-sigma events or excursions of a variable away from its mean).
    A test for which the underlying statistics are Gaussian.
    1: .158655; 2: .02275; 3: .00135

13. What is the t-test, based on the t-distribution? Find a table or Web page to answer the p-value question above, for a t-test with sample number N=10. How small is the difference from a Z-test?
    Any hypothesis test done when the data used fits a t-distribution.
    1: .1694 (~5% error); 2: .035402 (~55% error); 3: .00604 (~350% error)

14. What is the chi-squared distribution? When might you use it in a statistical test? (hint: where in science do we see a sum of squares involving just two or three variables?)
    Sum of squares of n number of normal, independent variables
    Correlations of variables that should be unrelated; common in weather forecasting

15. What is the F-test? Where have we seen a ratio of variances before? (hint: what was r-squared in linear regression)?
    A test done where the test statistic has an F-distribution, which is the null distribution in the analysis of variance.
    We've seen it in linear regression (regression error/total error in order to get best fit).

16. What are the joint PDF (probability density or distribution function)? marginal? conditional? Illustrate these 3 quantities with a sketch or annotation on an example like at https://en.wikipedia.org/wiki/Multivariate_normal_distribution.
    Multi-dimensional PDFs that cover multiple independent variables, each of which is Gaussian. Marginal is the PDF for only one variable within the set. Conditional is the distribution of one or more variables given a particular value for another. Pictured below is joint with marginal on the top/right axes, and conditional would be a "slice"/line through this graph.