For Questions 1-5, consider the School Admission dataset used in Homework 1.

<table>
<thead>
<tr>
<th></th>
<th>Accepted</th>
<th>Turned Away</th>
<th>Wait-listed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>110</td>
<td>133</td>
<td>49</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>485</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>White</td>
<td>336</td>
<td>359</td>
<td>251</td>
</tr>
</tbody>
</table>

Question 1. (2 points) These are the counts, which summarize the data. How would the raw data look like?

Question 2. (2 points) When these data are available, what is the main question that one is likely to ask? State that question and then formulate it as a hypothesis testing problem.

Question 3. (2 points) The Pearson’s statistic is a function of the observed counts and the (null) expected counts. Explain how you would compute the expected counts. Detail that for one of the (nine) cells. No need to simplify numerical expressions.

Question 4. (2 points) Define the permutation p-value for that statistic. Remember to define any math symbols you use. Why is it (typically) challenging to compute it?

Question 5. (4 points) Describe how this p-value is estimated in practice.
For Questions 6-10, assume that $X_1, \ldots, X_n$ are iid from a distribution $F$.

**Question 6.** (2 points) Define the empirical distribution (aka sample distribution) denoted $\hat{F}$. [Hint: You can do so by explaining how one samples from that distribution.]

**Question 7.** (2 points) Explain in a couple of sentences, using simple words, what the *bootstrap world* is. In particular, what corresponds to $\gamma$, defined as the 2nd moment of $F$, in the *bootstrap world*? Denote this by $\hat{\gamma}$.

**Question 8.** (4 points) Explain how you would compute a bootstrap 90% confidence interval based on $\hat{\gamma} - \gamma$ as pivot. (You do not need to know what $\hat{\gamma}$ is to answer this question.)

**Question 9.** (2 points) Describe the Kolmogorov-Smirnov (KS) test for testing $F = F_0$, where $F_0$ is some given distribution. Write down the test statistic.

**Question 10.** (4 points) Explain how you would obtain a p-value for the KS test statistic by Monte Carlo simulations.