Okay, here's my response to the question, formatted as if I were taking a graduate statistics exam:

The null hypothesis, 𝐻0: 𝐹(𝑧) = 𝐺(𝑧) in the context of the runs test, being equivalent to testing for random sampling stems from the underlying assumptions and properties of what constitutes a "random" sequence of observations. Here's a breakdown:

1. \*\*What Random Sampling Implies:\*\* Random sampling, in the context of the runs test (and often in general), means that:

\* \*\*Independence:\*\* Each observation is independent of the others. The outcome of one observation does \*not\* influence the outcome of any other observation.

\* \*\*Identical Distribution (Identically Distributed):\*\* All observations are drawn from the \*same\* underlying probability distribution. This means the data points, regardless of their order, share the same probabilistic characteristics defined by their cumulative distribution function (CDF).

2. \*\*The Role of CDFs (F(z) and G(z))\*\*:

\* \*F(z)\* and \*G(z)\* represent the \*cumulative\* distributions from which the two samples of data are obtained. If 𝐻0: 𝐹(𝑧) = 𝐺(𝑧), this means that the underlying probability distributions that generated the observations are identical. The values of \*z\* refer to the domain of the variable of interest and the values \*F(z)\* and \*G(z)\* are the respective probabilities. Therefore, whether an individual observation is of type F or G is probabilistically indistinguishable.

3. \*\*How the Runs Test Operates:\*\* The runs test focuses on the \*order\* or \*sequence\* of the data, specifically the number of "runs." A run is a sequence of consecutive identical observations. For example, in the sequence "FFGGFGGGF," there are five runs (F, GG, F, GGG, F). If the data are randomly sampled, we would expect the observations to be thoroughly mixed. The number of runs would not be notably high, nor notably low.

4. \*\*Connection to the Null Hypothesis:\*\* If the null hypothesis is true (the CDFs are equal), then there is no systematic difference between the types of observations (e.g. F and G) and they come from the same distribution. This is consistent with random sampling, where there is no systematic \*pattern\* or predictability in the sequence. If, however, the null hypothesis is false, the two observations types are dissimilar in some way. This would result in an \*unusual\* number of runs.

\* \*\*If F(z) and G(z) are \*different\*\*\*, then the data are not from the same distribution, meaning that their order can be predicted.

5. \*\*Non-Randomness and Runs:\*\* Deviations from random sampling manifest as:

\* \*\*Too few runs:\*\* This suggests \*clustering\* or \*grouping\* of observations of the same type (e.g., long stretches of "FF...F" followed by "GG...G"). This implies that the outcome of one observation is, in some way, related to the outcomes of the previous or subsequent observations, which would violate the assumption of independence.

\* \*\*Too many runs:\*\* This suggests excessive \*alternation\* or \*oscillation\* between the two types of observations, again violating independence and indicating some non-random structure. The number of runs will deviate from what one would expect under random mixing.

\*\*In conclusion:\*\* The equivalence between 𝐻0: 𝐹(𝑧) = 𝐺(𝑧) and testing for random sampling arises because equal CDFs directly support the assumptions of random sampling: identically distributed data and independence. When the data is not from the same distribution, this indicates the data is not being sampled randomly because the distribution can vary. The runs test detects deviations from these assumptions by assessing the sequence of the data to see if the number of runs is consistent with what would be expected under a purely random process. A significantly different number of runs indicates a non-random pattern, leading to rejection of the null hypothesis and, consequently, rejection of the assumption of random sampling.