C50 Enumerative & Asymptotic Combinatorics

Solution to Prize Question — Spring 2003

We use the result of Problem 15 on Problem Sheet 2, the identity

$$\sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k} = 2^{2n}.$$

Following the hint, we first calculate the expected number of times during the trial when the numbers of patients receiving drug and placebo are equal. This is obtained by summing, over all n-element subsets A of $\{1,\ldots,2n\}$, the number of values of k for which $|A \cap \{1,\ldots,2k\}| = k$, and dividing by the number $\binom{2n}{n}$ of subsets. Now the sum can be calculated by counting, for each value of k, the number of n-subsets A for which $|A \cap \{1,\ldots,2k\}| = k$, and summing the result over k.

For a given k, the number of subsets is $\binom{2k}{k}\binom{2(n-k)}{n-k}$, since we must choose k of the numbers $1,\ldots,2k$, and $n-k$ of the numbers $2k+1,\ldots,2n$. Hence, by the stated result, the sum is 2^{2n}, and the average is $2^{2n}/\binom{2n}{n}$.

Now consider the doctor's guesses in any particular trial. At any stage where equally many patients have received drug and placebo, he guesses at random, and is equally likely to be right as wrong. Such points contribute zero to the expected number of correct minus incorrect guesses. In each interval between two consecutive such stages, say $2k$, and $2l$, the doctor will guess right one more time than he guesses wrong. (For example, if the $2k$th patient gets the drug, then between stages $2k+1$ and $2l$ the number of patients getting the drug is $l-k-1$ and the number getting the placebo is $l-k$, but the doctor will always guess the placebo.) So the expected number of correct minus incorrect guesses is the number of such intervals, which is one less than the number of times that the numbers are equal.

So the expected number is $2^{2n}/\binom{2n}{n} - 1$, which is asymptotically $\sqrt{\pi n} - 1$, by Question 2 on Problem Sheet 2.