

Fisher's exact test

- Fisher's exact test is "exact" because it guarantees the α rate, regardless of the sample size
- Example, food additive and 14 mice

	Tumor	None	Total
Old	4	3	7
New	2	5	7
Total	6	8	

Fisher's exact test

Additive: 0 0 0 0 0 0 0 0 N N N N N N N
Tumor : T T T T F F F T T F F F F F F

permuted

Additive: 0 0 0 0 0 0 0 0 N N N N N N N
Tumor : T T T F F T F T F F F F T F

permuted again

Additive: 0 0 0 0 0 0 0 0 N N N N N N N
Tumor : T F F F T T F F F T T F T F

Hyper-geometric distribution

- X number of tumors for the old additive
- Y number of tumors for the new additive
- $H_0 : p_1 = p_1 = p$
- Under H_0

Proof

$$P(X = x) = \binom{n_1}{x} p^x (1 - p)^{n_1 - x}$$

$$P(Y = z - x) = \binom{n_2}{z - x} p^{z - x} (1 - p)^{n_2 - z + x}$$

$$P(X + Y = z) = \binom{n_1 + n_2}{z} p^z (1 - p)^{n_1 + n_2 - z}$$

Fisher's exact test

- More tumors under the old than the new additive
- Calculate an *exact* P-value
- Use the conditional distribution = hypergeometric
- Fixes both the row and the column totals
- Yields the same test regardless of whether the rows or columns are fixed
- Hypergeometric distribution is the same as the permutation distribution given before

Tables supporting H_a

- Consider $H_a : p_1 > p_2$
- P-value requires tables as extreme or more extreme (under H_a) than the one observed
- Recall we are fixing the row and column totals

Calculations

$$\begin{aligned} P(\text{Table 1}) &= P(X = 4 | X + Y = 6) \\ &= \frac{\binom{7}{4} \binom{7}{2}}{\binom{14}{6}} = .245 \end{aligned}$$

Notes

- Two sided p-value = $2 \times$ one sided P-value
(There are other methods which we will not discuss)
- P-values are usually large for small n
- Doesn't distinguish between rows or columns
-

Monte Carlo

Observed table $X = 4$

```
Additive: 0 0 0 0 0 0 0 0 N N N N N N N
Tumor      : T T T T F F F T T T F F F F
```

Permute the second row

```
Additive: 0 0 0 0 0 0 0 0 N N N N N N N
Tumor      : T T F F F T F T T F F F T F
```

Simulated table $X = 3$

Do over and over

Calculate the proportion of tables for which
the simulated $X \geq 4$

This proportion is a Monte Carlo estimate
for Fisher's exact P-value