Polya Sampling Scheme Take $w \geq 1$, $b \geq 1$.

A) The ~~total~~ no. of possible sequences can be obtained by noting that there are

1) $w+b$ choices for the first ball; THEN
2) $w+b+1$ .. .. .. second .. .. ..
⋮
n) $w+b+n-1$ .. .. .. n-th ball.

Using the multiplication principle, there are thus

$$(w+b)(w+b+1)\cdots(w+b+n-1) = \frac{(w+b+n-1)!}{(w+b-1)!}$$

[~~Even if $w+b=1$, this still makes sense if we take $0!=1$~~]

B) 1) Regardless of where in the sequence the x w's occur (take $1 \leq x \leq n-1$ for now), there are

1) w choices for the first, white ball, then
2) $w+1$ .. .. .. second, white ball
⋮
x) $w+x-1$ .. .. .. x-th .. ..

So there are $w(w+1)\cdots(w+x-1) = \dfrac{(w+x-1)!}{(w-1)!}$

(which makes sense even for $1 = w$ if $0! = 1$)

possible ways to get a sequence of x wh balls ($x \geq 1$) under the Polya sampling scheme.

2) Similarly there are (for $x \leq n-1$, i.e. $n-x \geq 1$),

1) b choices for 1st black

2) $b+1$ . . . 2nd black

⋮

$n-x$) $b+n-x-1$ . . . $(n-x)$th black

So there are $b(b+1)\cdots(b+n-x-1) = \dfrac{(b+n-x-1)!}{(b-1)!}$

(which makes sense even if $b=1$, with $0!=1$).

~~So there are~~ [NOTE: the "edge" cases $x=n$ and $x=0$ give $\dfrac{(w+n-1)!}{(w-1)!}\cdot 1$ and $1\cdot \cancel{(b+n-x-1)!}\,\dfrac{(b+n-1)!}{(b-1)!}$ respectively]

---

thus, according to the prescription (A) from Lecture 6,

$$P(\text{exactly } x \text{ wh}) = \frac{\dfrac{n!}{x!(n-x)!}\cdot\dfrac{(w+x-1)!}{(w-1)!}\cdot\dfrac{(b+n-x-1)!}{(b-1)!}}{\dfrac{(w+b+n-1)!}{(w+b-1)!}}$$

$$= \frac{\dfrac{(w-1+x)!}{(w-1)!\,x!}\;\dfrac{(b-1+n-x)!}{(b-1)!\,(n-x)!}}{\dfrac{(w+b-1+n)!}{(w+b-1)!\,n!}} = \frac{\binom{w-1+x}{x}\binom{b-1+n-x}{n-x}}{\binom{w+b-1+n}{n}}$$

this is true for all $x = 0, 1, \ldots, n$ (remember, taking $0! = 1$, further taking $\binom{0}{0} = \frac{0!}{0!\,0!} = 1$)

This ~~eqn~~ ran is a BETA-BINOMIAL distribution

This is (most likely) a NEW prob dist'n for most. It can be used for modelling the no. of boys (or girls) in a family of a given SIZE (n children) and tends to fit (across many families) much better than the binomial. We shall meet this dist'n again later in the course when we meet MIXTURE MODELS.

Examples

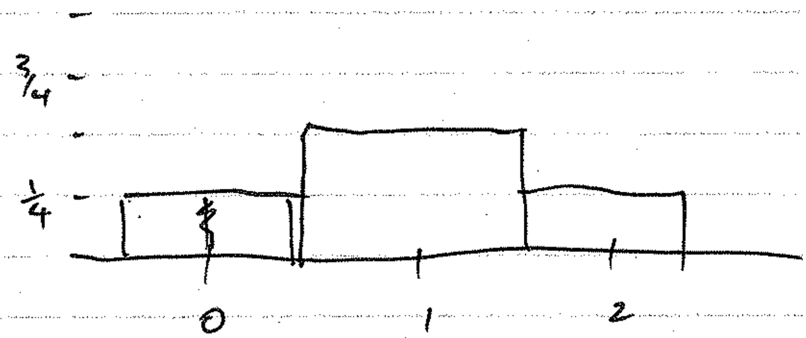
Consider all 3 schemes where $b = w = 1$, $n = 2$.

1) Sampling with Replacement.

$$P(x \text{ wh}) = \binom{2}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{2-x} = \binom{2}{x}\Big/4$$

$x = 0, 1, 2.$

$P(0 \text{ wh}) = \frac{1}{4}$
$P(1 \text{ wh}) = \frac{1}{2}$
$P(2 \text{ wh}) = \frac{1}{4}$

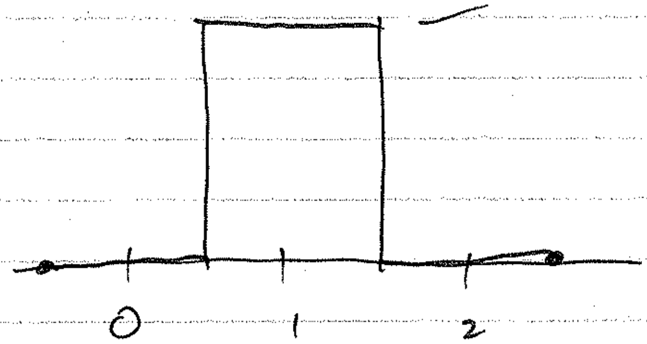


2) Sampling w/o Repl.

$$P(x \text{ wh}) = \frac{\binom{1}{x}\binom{1}{2-x}}{\binom{2}{2}}$$

BUT!! remember for both binomial coeffs to be well defined $x \le 1$ and $2-x \le 1$ i.e. $x \ge 1$

So $P(0 \text{ wh}) = 0$
$P(1 \text{ wh}) = 1$
$P(2 \text{ wh}) = 0$

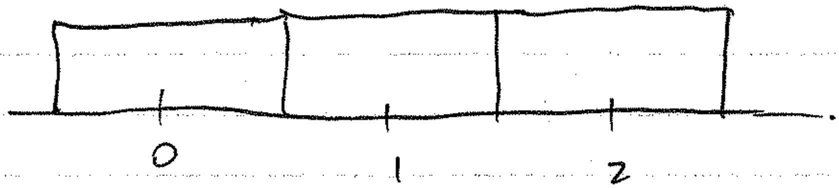


3) Polya Scheme.

$$P(x\ \text{wh}) = \frac{\binom{x}{x}\binom{2-x}{2-x}}{\binom{3}{2}} \qquad x = 0, 1, 2$$

$$P(0\ \text{wh}) = \frac{1}{3}$$

$$P(1\ \text{wh}) = \frac{1}{3} \qquad P(2\ \text{wh}) = \frac{1}{3}$$



(Discrete Uniform)