Power of a Binomial Test

T.L. Scofield

9/23/2015

Rejection region for 100 coin flips

From the command

```
qbinom(.025, 100, .5)
```

```
## [1] 40
```

we learn that, for a binomial random variable $X \sim \text{Binom}(100, .5)$, the cumulative probability up to but not including $X = 40$ is 0.025. Actually, that is not quite true, since

```
pbinom(40, 100, .5)
```

```
## [1] 0.02844397
```

```
pbinom(39, 100, .5)
```

```
## [1] 0.0176001
```

which shows $P(X \leq 39) = 0.0176$, while $P(X \leq 40) = 0.0284$; we cannot hit 0.025 exactly. Since, the two-tailed area

$$P(X \leq 39 \,\text{or}\, X \geq 61) = 2 \cdot P(X \leq 39) \doteq 0.0352,$$

while the two-tailed area

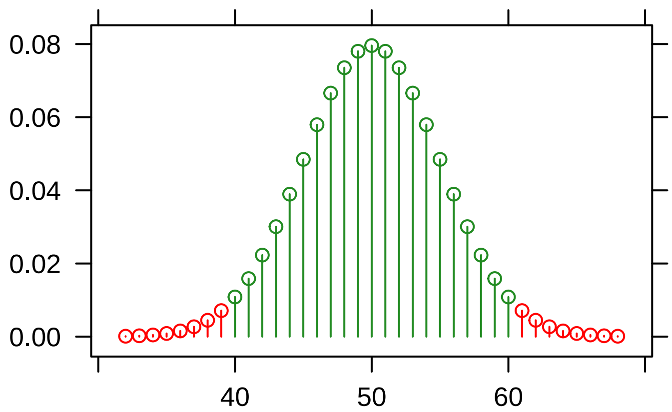
$$P(X \leq 40 \,\text{or}\, X \geq 60) = 2 \cdot P(X \leq 39) \doteq 0.05688,$$

the former is the appropriate **rejection region** for an hypothesis test involving the count of heads in 100 flips from a coin with hypotheses

$$H_0 \colon \pi = 0.5, \qquad H_a \colon \pi \neq 0.5$$

and significance level $\alpha = 0.05$. We display the null distribution along with the rejection region in red:

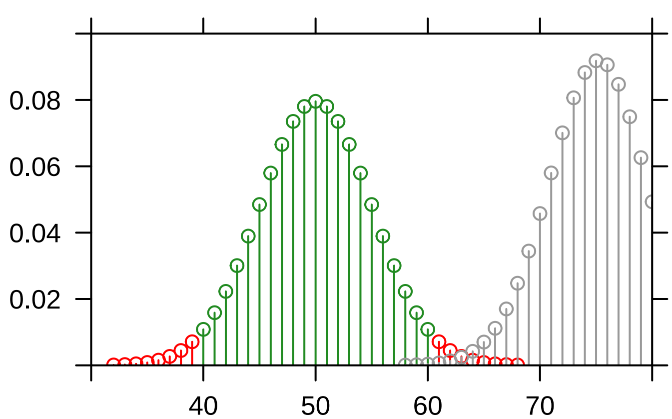
```
plotDist("binom", params=c(100, .5), col=c("red","forestgreen"),
         groups=abs(x-50) <= 10)
```



Computing β, the probability of Type II Error

Suppose our coin actually has a probability of landing "heads" equaling 0.75. Then, counter to what is hypothesized in H_0, $X \sim \text{Binom}(100, 0.75)$. We overlay this distribution (displayed in gray) with the null distribution.

```
plotDist("binom", params=c(100, .5), col=c("red","forestgreen"),
         groups=abs(x-50) <= 10, xlim=c(30,80), ylim=c(0,0.1))
plotDist("binom", params=c(100, .75), col="gray60", add=TRUE)
```



The probability of making a Type II error, β, should be small, as the likelihood of values from our coin (with $\pi_a = 0.75$) falling in the green region (where the null hypothesis is *not* rejected) appears to be small. We can find its actual value with commands like

```
sum(dbinom(40:60, 100, 0.75))
```

```
## [1] 0.0006865922
```

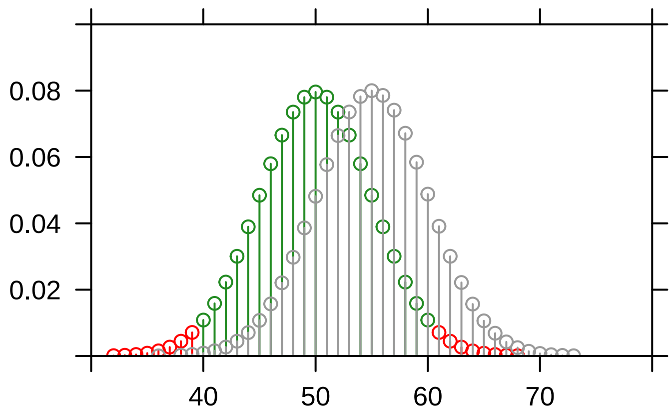
or

```
pbinom(60, 100, .75) - pbinom(39, 100, .75)
```

```
## [1] 0.0006865922
```

Now, if our coin has a probability of "heads" equaling 0.55, the likelihood of Type II error should rise. The gray distribution (corresponding to how the coin actually behaves) has a lot more of its probability lying inside the nonrejection region.

```
plotDist("binom", params=c(100, .5), col=c("red","forestgreen"),
         groups=abs(x-50) <= 10, xlim=c(30,80), ylim=c(0,0.1))
plotDist("binom", params=c(100, .55), col="gray60", add=TRUE)
```



We compute β as before, seeing (as predicted) it is much larger than before.

```
pbinom(60, 100, .55) - pbinom(39, 100, .55)
```

```
## [1] 0.8648077
```

So, β can only be calculated when we make a presumption about π_a, the probability our coin produces a "head". Not only does its value depend on how far away the true value of π is from what is hypothesized, but it also depends on the choice of significance level α.

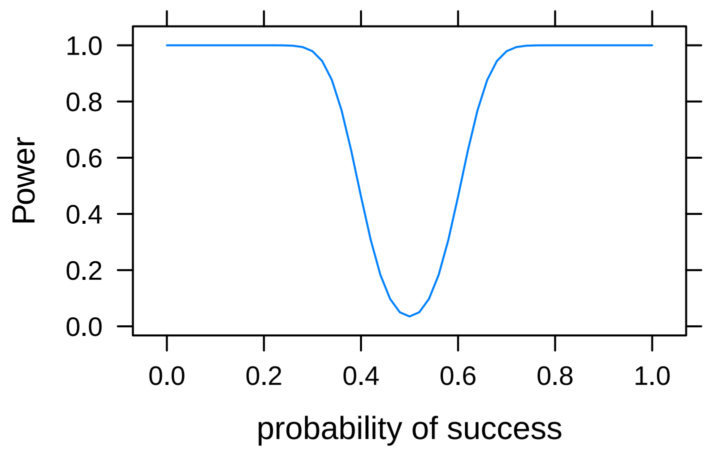
Power

Power is defined as the probability a false null hypothesis is rejected. So

$$\text{power} \ = \ 1 - P(\text{not rejecting a false } H_0) \ = \ 1 - \beta.$$

Like β, it relies on α and knowledge of π_a, making it difficult to calculate. We may illustrate how the power of a binomial test changes as π_a changes.

```
piAlt = seq(0, 1, .02)
myBeta = pbinom(60, 100, piAlt) - pbinom(39, 100, piAlt)
xyplot(1-myBeta ~ piAlt, type="l", xlab="probability of success", ylab="Power")
```



You can, in fact, increase the power of a binomial test at any fixed value of π_a and α by increasing the sample size n. Our next plot gives power for different choices of n, assuming that $\pi_a = 0.55$ and $\alpha = 0.05$.

```
enn = 1:2000
critical = qbinom(.025, enn, .5)
beta = pbinom(enn-critical,enn,.55) - pbinom(critical-1,enn,.55)
xyplot(1-beta ~ enn, type="l", lwd=0.5, xlab="n", ylab="power")
```