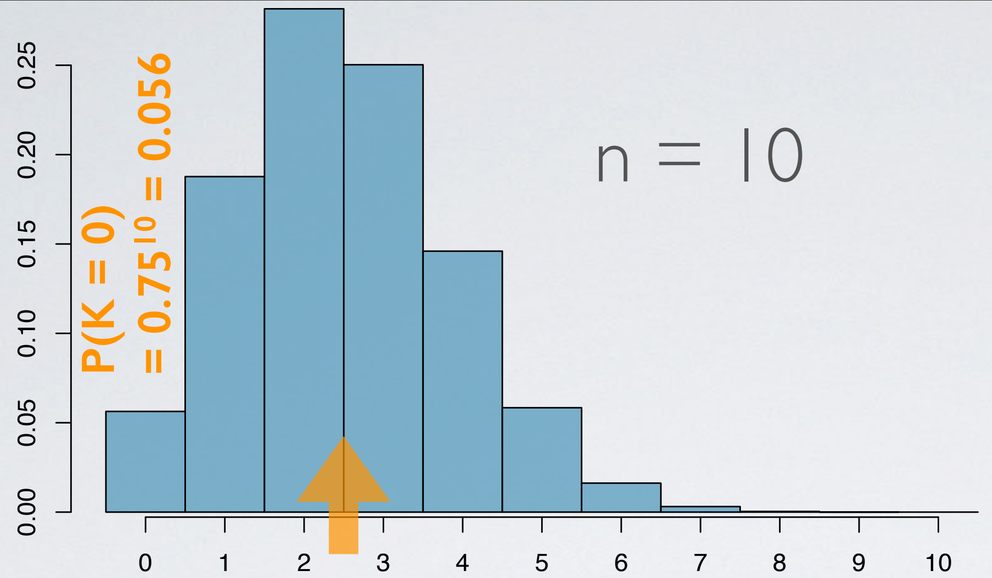
normal approximation to binomial

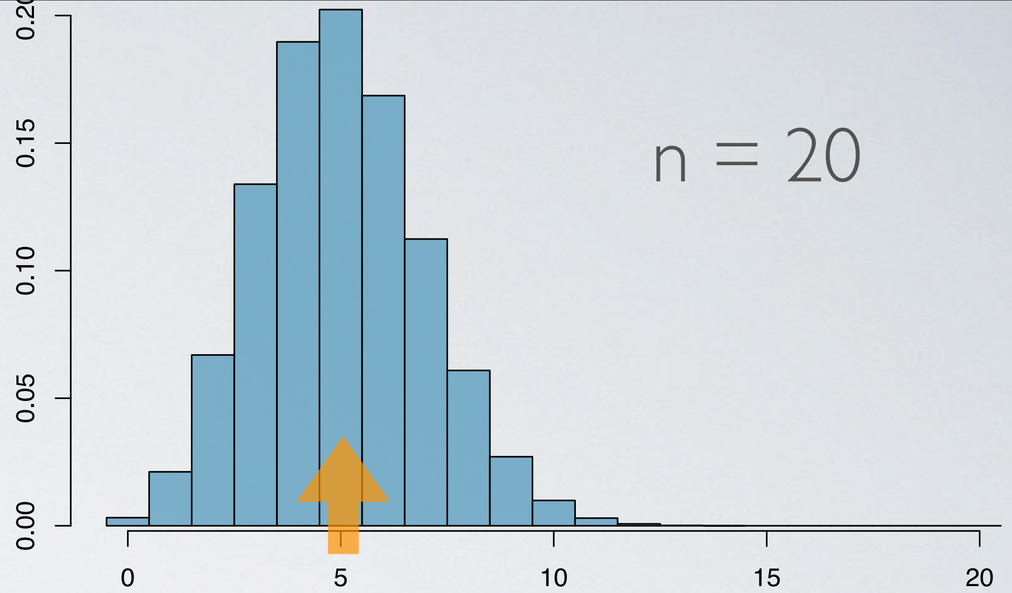
- shapes of binomial distributions
- normal approximation

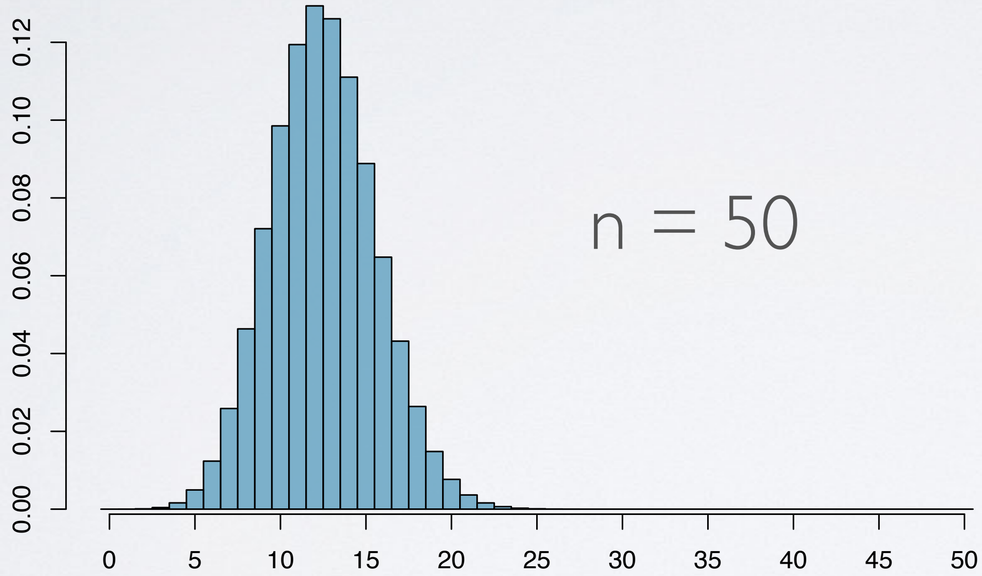
Duke UNIVERSITY

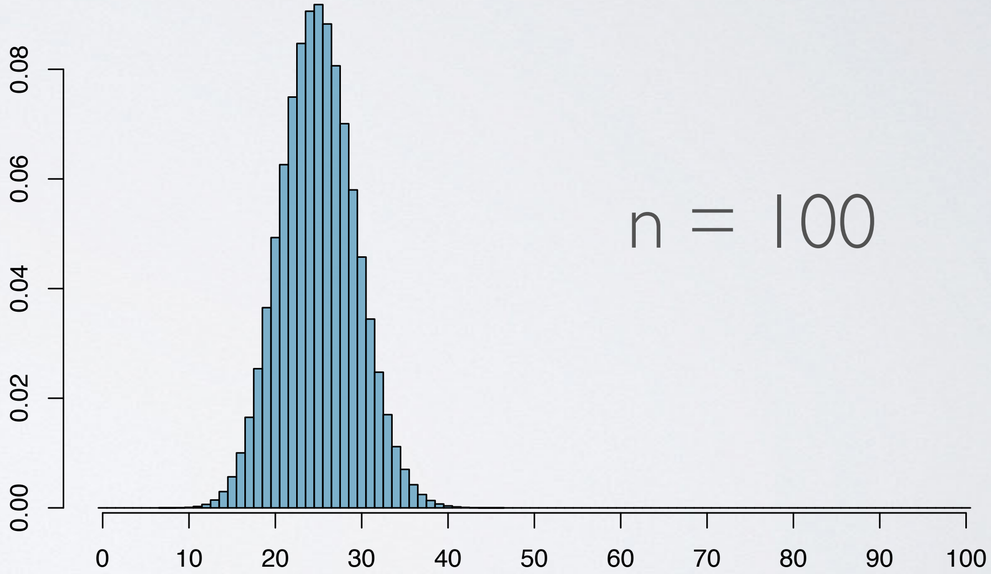
Dr. Mine Çetinkaya-Rundel
Duke University

p = 0.25









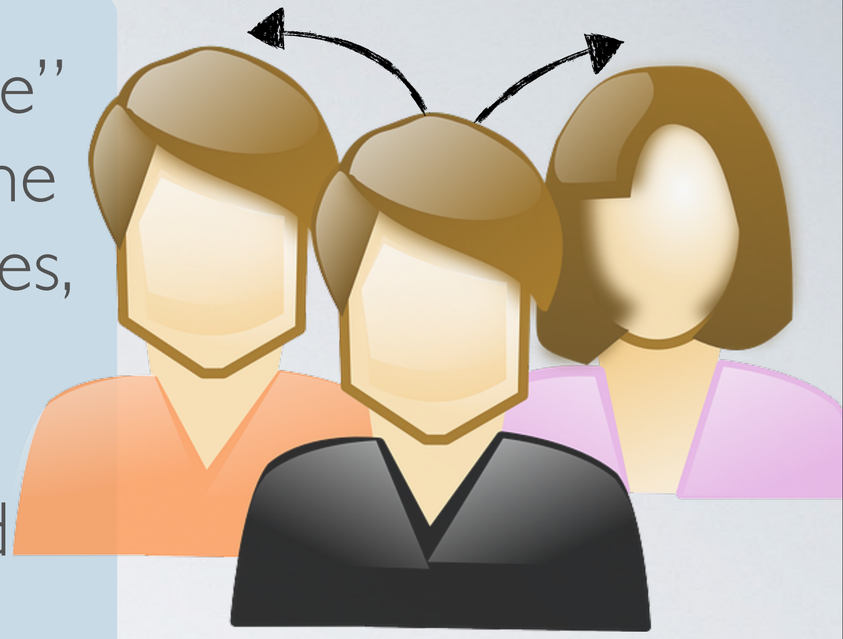
Recent study: "Facebook users get more than they give"

- friend requests: 40% made, 63% received at least one
- likes: liked 14 times, had their content "liked" 20 times, on average
- messages: sent 9 messages, received 12, on average
- tags: 12% tagged a friend in a photo, but 35% tagged

other findings:

- 25% considered power users
- average Facebook user has 245 friends

P(70 or more power user friends) = ?



$p = 0.25$

$n = 245$

$P(K \geq 70) = ?$

Image: http://pixabay.com/p-35681/?no_redirect

Poll: http://www.pewinternet.org/Reports/2012/Facebook-users/Summary.aspx

$P(K \geq 70) = ?$

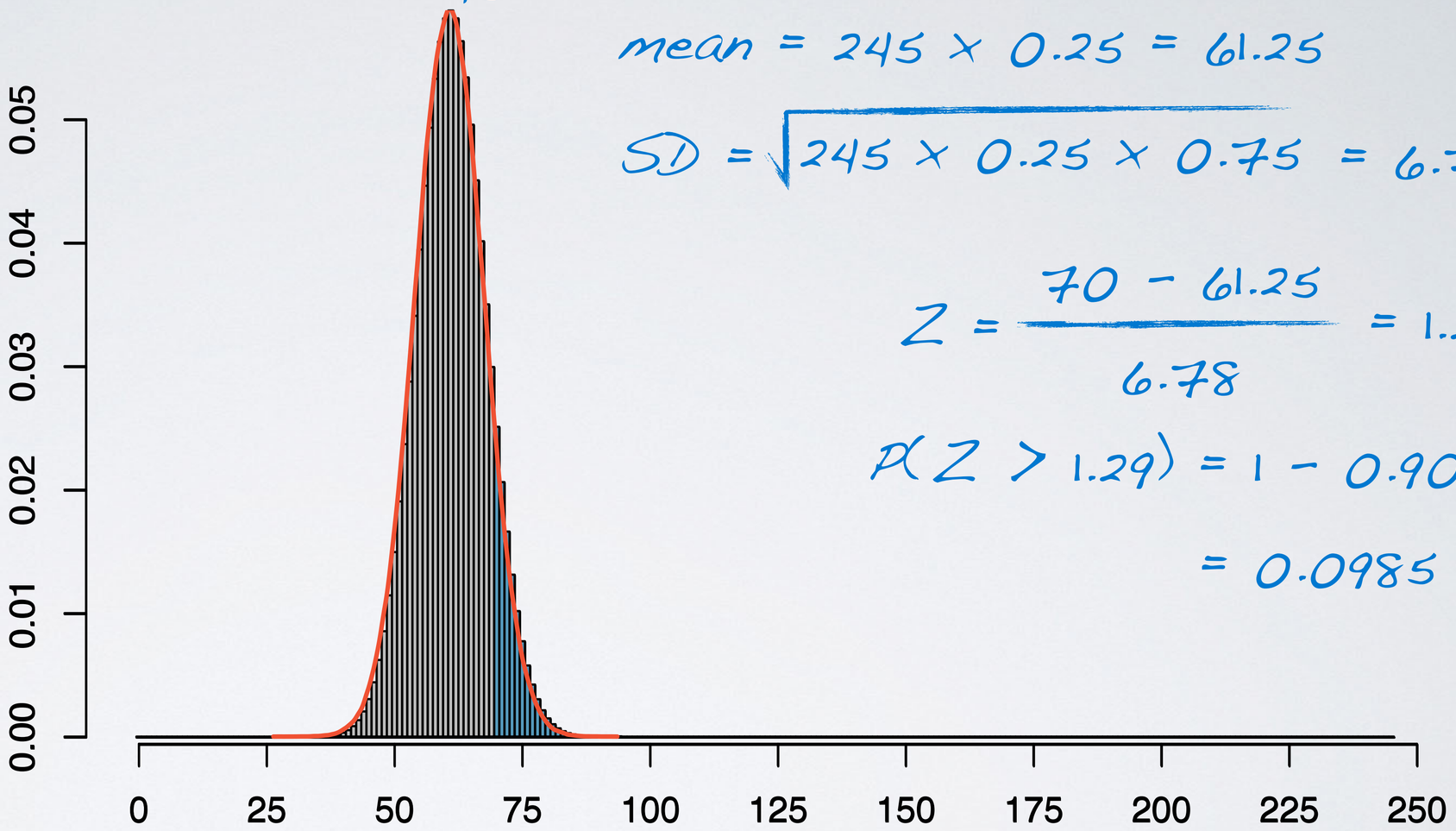
$= P(K = 70) + P(K = 71) + \ldots + P(K = 245)$

(1) ✓ $n = 245$, fixed

(2) ✓ power user / not

(3) ✓ $p = 0.25$

(4) ~ independence



$N(\text{mean}, SD)$

$\text{mean} = 245 \times 0.25 = 61.25$

$SD = \sqrt{245 \times 0.25 \times 0.75} = 6.78$

$$Z = \frac{70 - 61.25}{6.78} = 1.29$$

$P(Z > 1.29) = 1 - 0.9015$

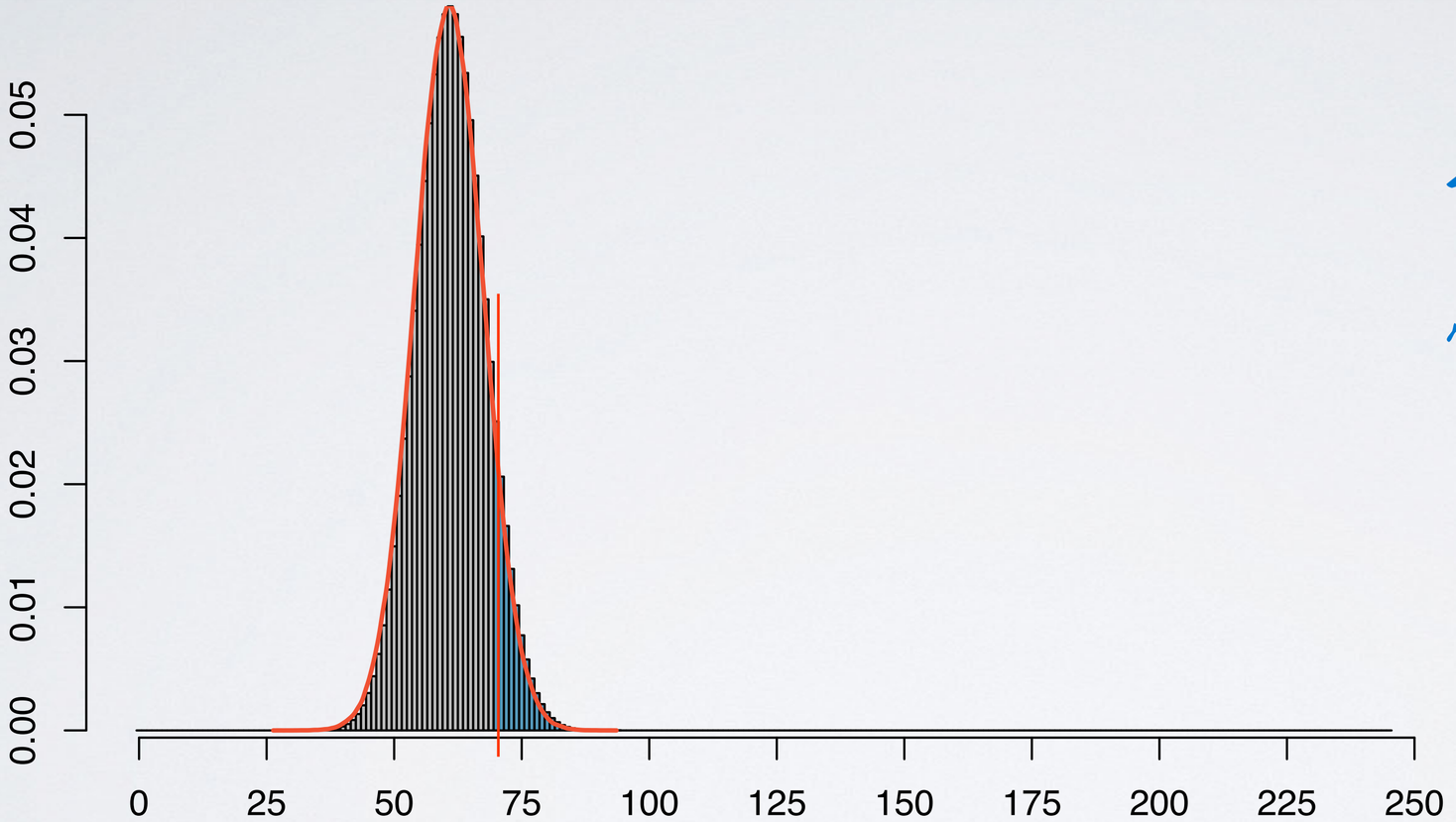
$= 0.0985$

Z	Second decimal place of Z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015

R

```
> sum(dbinom(70:245, size = 245, p = 0.25))
[1] 0.113
```

vs 0.0985???

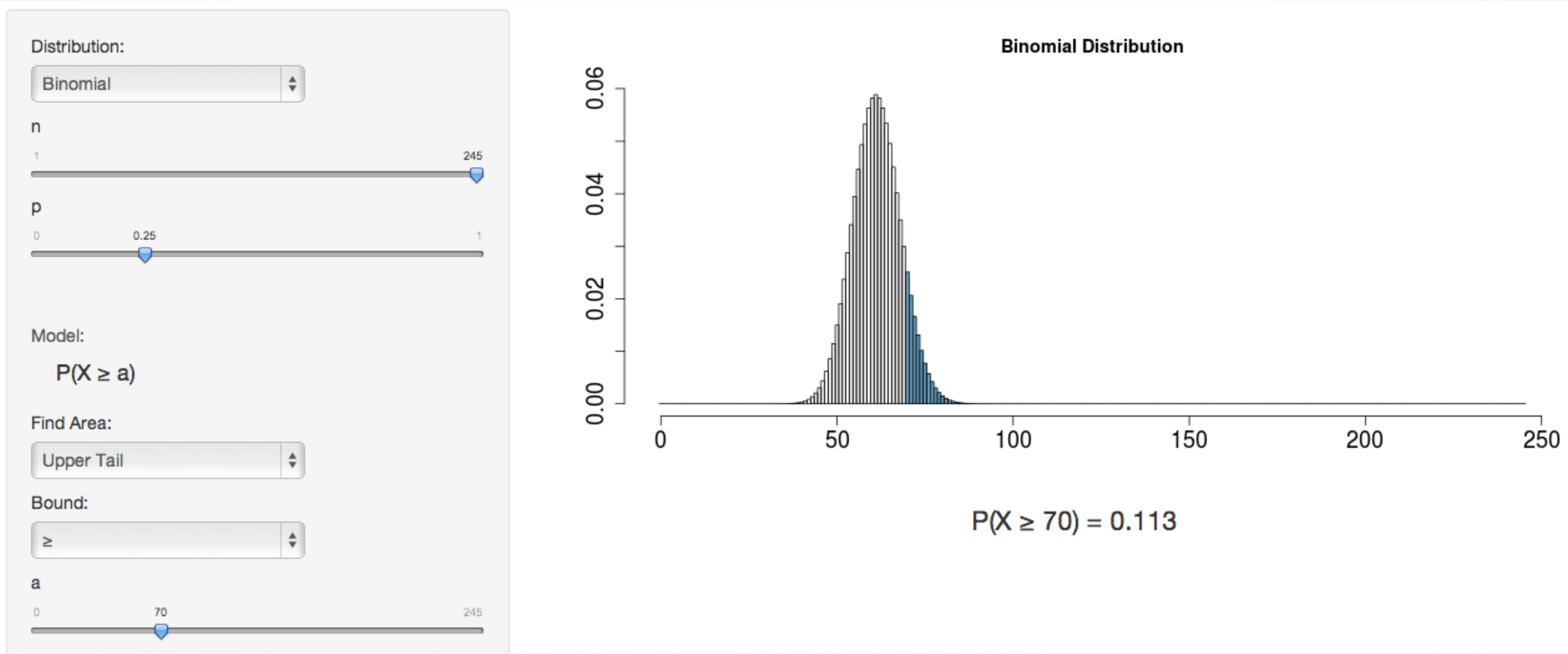


$$Z = \frac{69.5 - 61.25}{6.78} = 1.22$$

$$P(Z > 1.22) = 1 - 0.8888$$

$$= 0.1112$$

http://bit.ly/dist_calc



Success-failure rule: A binomial distribution with at least 10 expected successes and 10 expected failures closely follows a normal distribution.

$$np \geq 10$$
$$n(1-p) \geq 10$$

Normal approximation to the binomial: If the success-failure condition holds,

$$\text{Binomial}(n,p) \sim \text{Normal}(\mu,\sigma)$$

where $\mu = np$ and $\sigma = \sqrt{np(1-p)}$

practice

What is the minimum required n for a binomial distribution with $p = 0.25$ to closely follow a normal distribution?

$n \times 0.25 \geq 10$

$n \geq 10 / 0.25$

$n \geq 40$

$n \times 0.75 \geq 10$

$n \geq 10 / 0.75$

$n \geq 13.33$