Example Problem: HIV prevalence in South Africa

According to UN AIDS*, HIV prevalence in South Africa was 17.8% among adults 15 to 49 years old in 2009. Assume this prevalence estimate is accurate today, and we randomly sample 500 individuals in South Africa. Suppose X is the number of HIV positive individuals in the sample.

Model X using the binomial distribution.

1. How many individuals do we expect to be HIV positive in the sample.

E(X) = np = 500*0.178 = 89

2. What is the standard deviation of the number of HIV positive individuals in the sample?

sd(X) = √np(1-p) = 8.553245

3. What is the probability of observing more than 100 HIV positive individuals?

```
. di 1 - binomial(500, 100, 0.178)

.09089616

. di binomialtail(500, 101, 0.178)

.09089616
```

4. What is the probability of observing between 85 and 95 HIV positive individuals?

```
. di binomial(500, 95, 0.178) - binomial(500, 84, 0.178)

.47533949

. di binomialtail(500, 85, 0.178) - binomialtail(500, 96, 0.178)

.47533949
```

Now, model X using the normal distribution instead.

1. What is E(X)?

E(X) = np = 500*0.178 = 89

2. What is sd(X)?

sd(X) = √np(1-p) = 8.553245

3. What is the probability of observing more than 100 HIV positive individuals?

P(X> 100) = P(Z > (100 – 89)/ 8.55) = P(Z > 1.286)

```
. di 1-normal(1.286)
.09922153
```

4. What is the probability of observing between 85 and 95 HIV positive individuals?

$P(85 < X < 95)$

$= P(X < 95) – P(X < 85)$

$= P(Z < (95 – 89)/ 8.55) – P(Z<(85-89)/ 8.55)$

$= P(Z< .702) – P(Z<-.468)$

```
. di normal(0.702) - normal(-0.468)
.43876812
```

5. Do the normal and binomial models give similar results?

What is the probability of observing more than 100 HIV positive individuals?

Binomial: .09089616

Normal: .09922153

What is the probability of observing between 85 and 95 HIV positive individuals?

Binomial: .47533949

Normal: .43876812

Yes, they give similar results. Approximation is better "in the tails", i.e. for calculating the probability of observing more than 100 HIV+ individuals; than in the center of the distribution (between 85 and 95 HIV+).

*http://www.unaids.org/en/regionscountries/countries/southafrica/