Stat310

Sequences of rvs

Hadley Wickham

1. Special events
2. Sequences
3. Chebyshev’s theorem
4. The law of large numbers

Guest lecture

How a Lack of Understanding of a 'Simple' Statistical Technique Causes Huge Losses and Poor Decisions in Sports (and Drug Development!) — **Scott Berry**

Monday Mar 19, 4-5pm
McMurtry Auditorium

(5 points extra credit)

Math background

Your turn

Simplify the following expressions (Hint: if you're stuck, think about n = 2)

$$\prod_{x=1}^{n} x \qquad \prod_{i=1}^{n} e^{x_i} \qquad \ln\left(\prod_{i=1}^{n} x_i\right)$$

Your turn

If X and Y are independent, what does E(XY) equal?

Sequences

Sequences

1 variable: X

2 variables: X, Y

...

n variables: $X_1, X_2, X_3, \ldots, X_n$

Why?

Random experiment

"A random experiment is an experiment, trial, or observation that can be **repeated numerous times under the same conditions**... It must in no way be affected by any previous outcome and cannot be predicted with certainty." (http://cnx.org/content/m13470/latest/)

Sequences

$X_i \sim \text{Normal}(\mu_i, \sigma_i)$

$X_i \sim \text{Normal}(\mu, \sigma_i)$

$X_i \sim \text{Normal}(\mu_i, \sigma)$

$X_i \sim \text{Normal}(\mu, \sigma)$

In this class, almost always assume that the X_i's are **independent**. In the last case they are also **identically distributed**.

iid = independent & identically distributed

Your turn

X_i are iid $N(0, 2)$.

What is $E(X_{30})$? What is $Var(X_{2001})$?

What is $Cor(X_{10}, X_{11})$? $Cor(X_1, X_{1000})$?

$$E(\sum_{i}^{n} X_i) = \sum_{i}^{n} E(X_i)$$

$$Var(\sum_{i}^{n} a_i X_i) = \sum_{i}^{n} a_i^2 Var(X_i)$$

If what is true?

$$E(\prod_{i}^{n} X_i) = \prod_{i}^{n} E(X_i)$$

If what is true?

Limits

Typically will define some function of n random variables, e.g. $\bar{X}_n$

What happens to $\bar{X}_n$ when n → ∞?

Why? Because often it will converge, and we can use this to approximate results for any large n. So we can solve a bigger class of problems!

Chebyshev

No limit - but a good starting point

$$P(|X - \mu| < K\sigma) \geq 1 - \frac{1}{K^2}$$

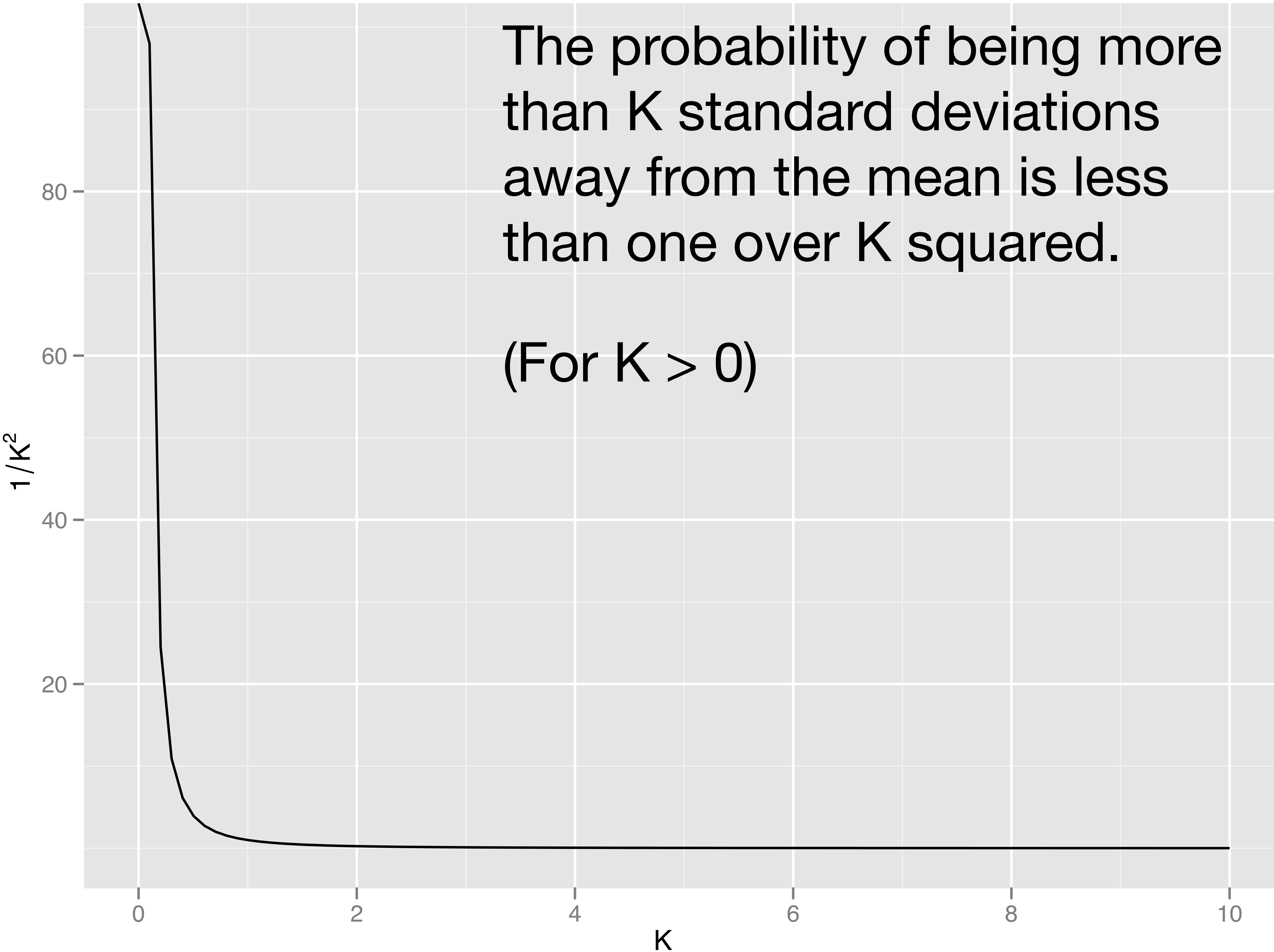
$$P(|X - \mu| > K\sigma) \leq \frac{1}{K^2}$$

For K > 0

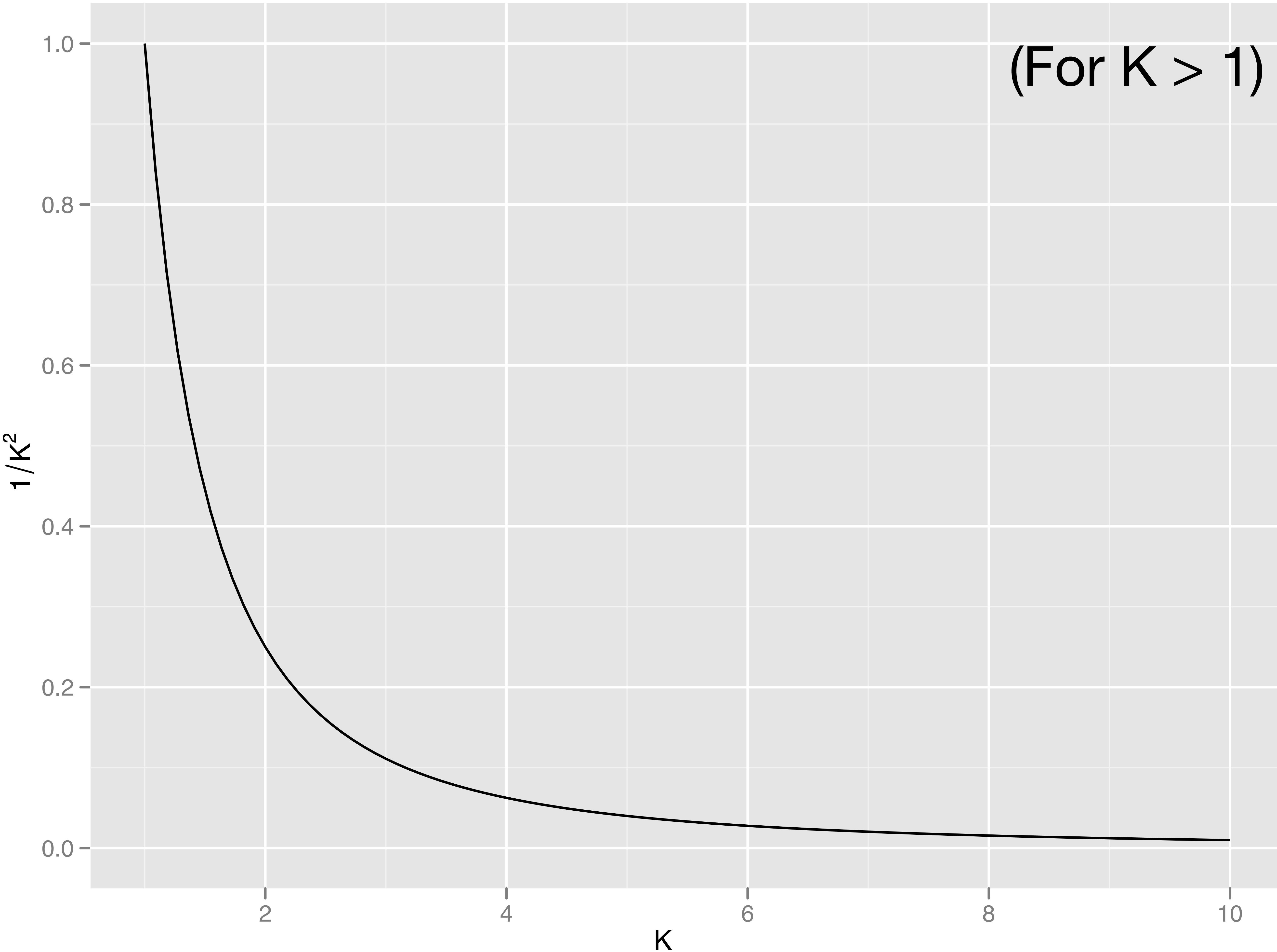
Your turn

How can you put this in words?

$$P(|X - \mu| > K\sigma) \leq \frac{1}{K^2}$$

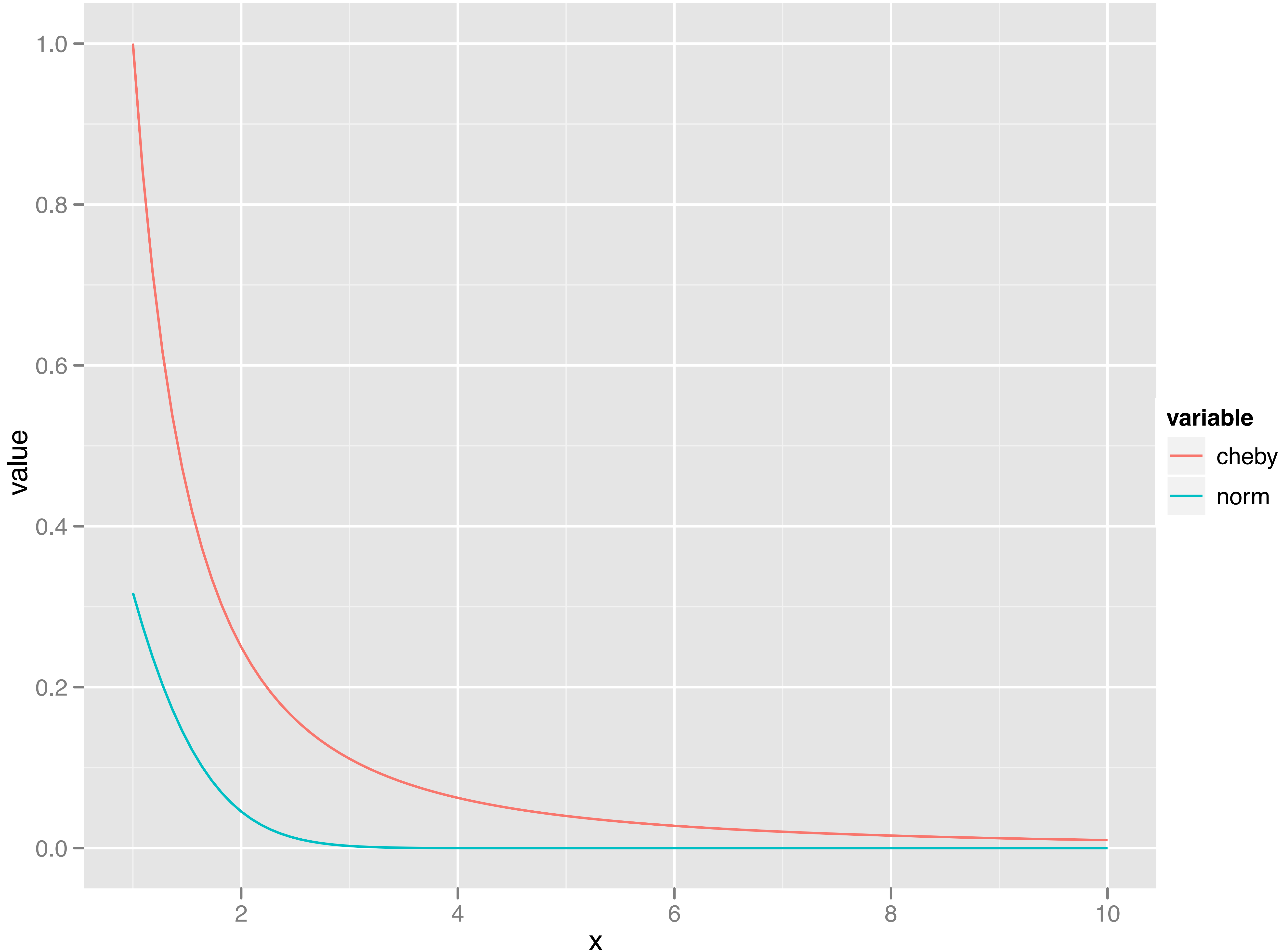


Tuesday, March 13, 12


(For K > 1)
1.0
0.8
0.6
0.4
0.2
0.0
$1/K^2$
2
4
6
8
10
K

Your turn

How does this compare to the normal distribution? Compare the probability of being less than 1, 2 and 3 standard deviations away from the mean given by Chebychev and what we know about the normal.



1.0
0.8
0.6
0.4
0.2
0.0
value
2
4
6
8
10
x
variable
cheby
norm

Law of large numbers

LLN

Law of large numbers

$X_1, X_2, ..., X_n$ iid. with $\mu, \sigma^2 < \infty$

There are four ways to write the result.

What does it mean?

As we collect more and more data, the sample mean gets closer and closer to the true mean.

Not that surprising!

But note that we didn't make any assumptions about the distributions

Example

You have a loaded die that has probability p of rolling a six, and $(1 - p)/5$ of rolling any other number.

What sequence of random experiments could you use to find p?