Dynamic Subclass Instantiation Distribution

Amy Wheeler

David Binkley

Loyola University Maryland
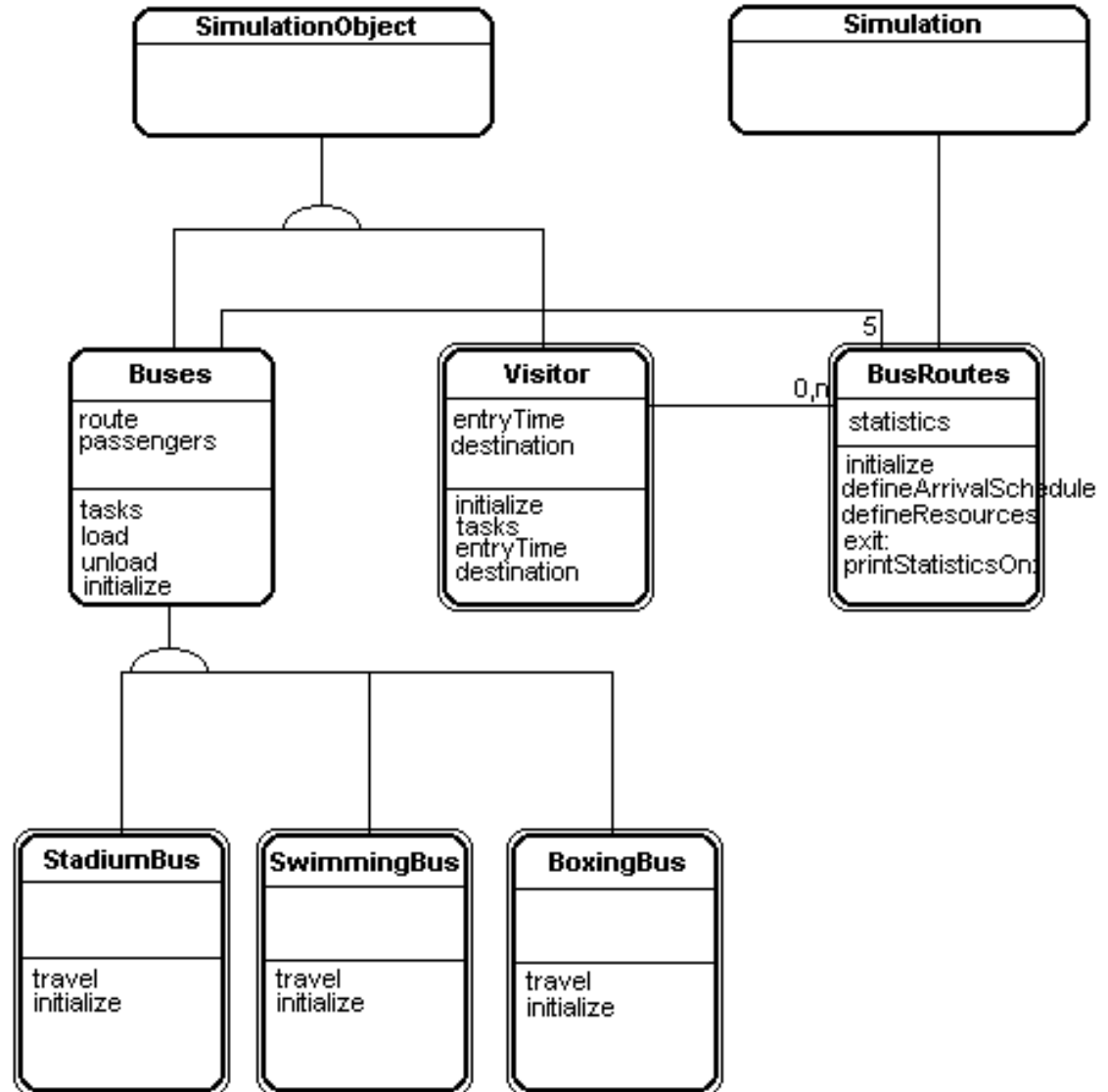
1 floor
3 full time faculty
0 PhD students
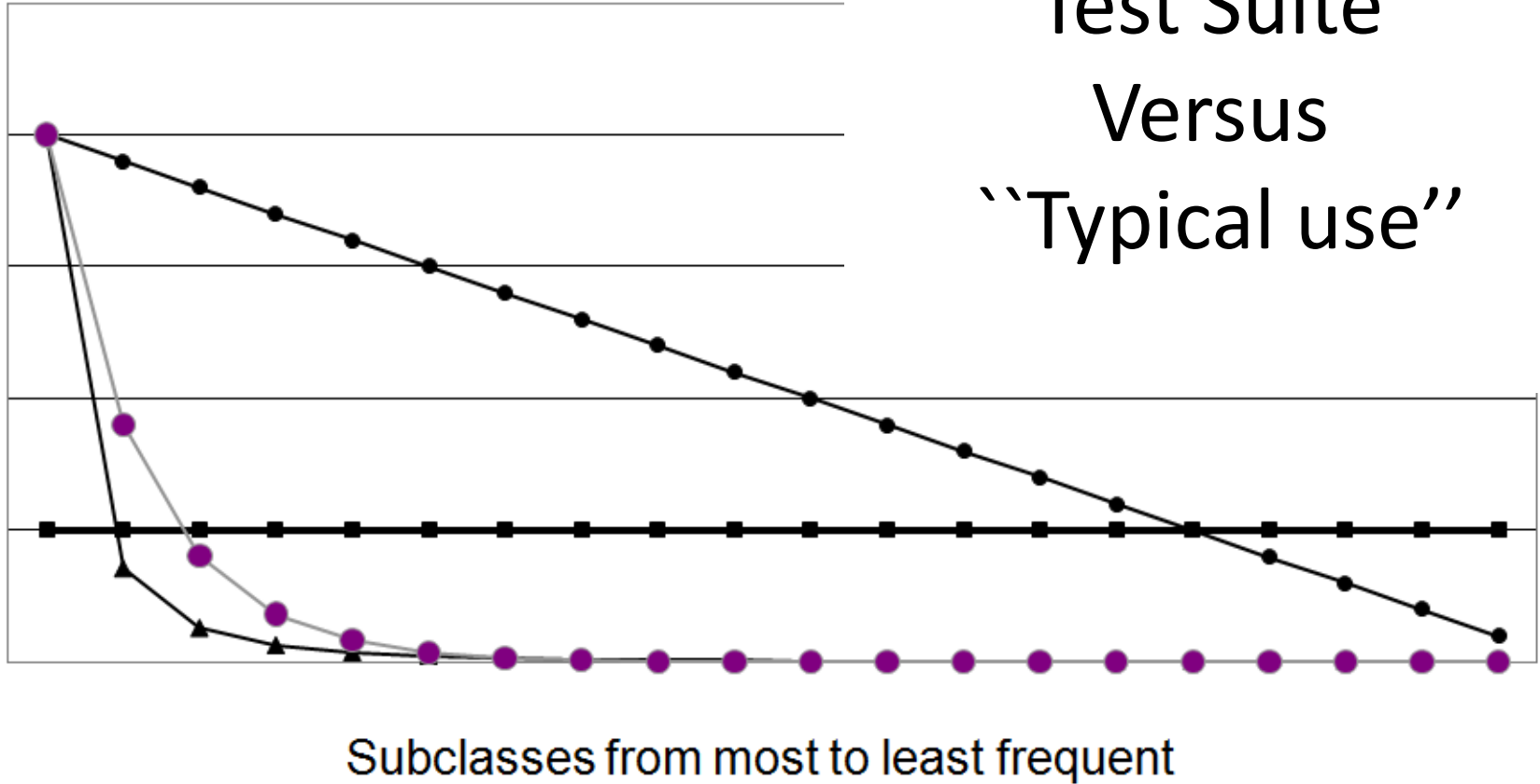
What

Dynamic subclass
instantiation
tracking

Example Distributions

Test Suite Versus ``Typical use''

Subclasses from most to least frequent

How: JBoss

Part 1: Join Point Description

```
<bind pointcut = "construction(*->new(..))"
        and !construction(Tracer*->new(..))" >
  <before aspect="Tracer"
            name="ConstructorInterceptor" />
</bind>
```

Part 2: Aspect

```
public Object Tracer::ConstructorInterceptor
    (ConstructionInvocation ci)
 {
   Class c = ci.getConstructor().getDeclaringClass();
   Chain invocation = new Chain();
   while (c != null)
   {
     invocation.addClass(c);
     c = c.getSuperclass();
   }
   return ci.invokeNext();
 }
```

Who – Subjects Programs

Program	LoC	Classes	Description
GanttProject 2.0.9	69K	564	Project Scheduling
jasmin 2.3	40K	216	Java Assembler
jess	no src	460	Sandia Rule Engine
jmeter 2.3.4	147K	792	Testing Tool
jolden	6215	20	Olden Bench Mark
jtopas	24K	65	Java Tokenizer Lib
nanoxml 1.2	95K	611	Java Parser
siena 0.9	98K	34	Event Services
Total	479K	2762	

Research Questions

1. **More than 1?**

   Do classes with instantiated subclasses have more than one instantiated subclass?

2. **Uniform or Skewed?**

   If multiple subclasses are instantiated, is the distribution of subclass instantiation uniform or skewed?

3. **Dominated? Correlated?**

   If the distribution is skewed, then does a
   a) single or do a few subclasses dominate?
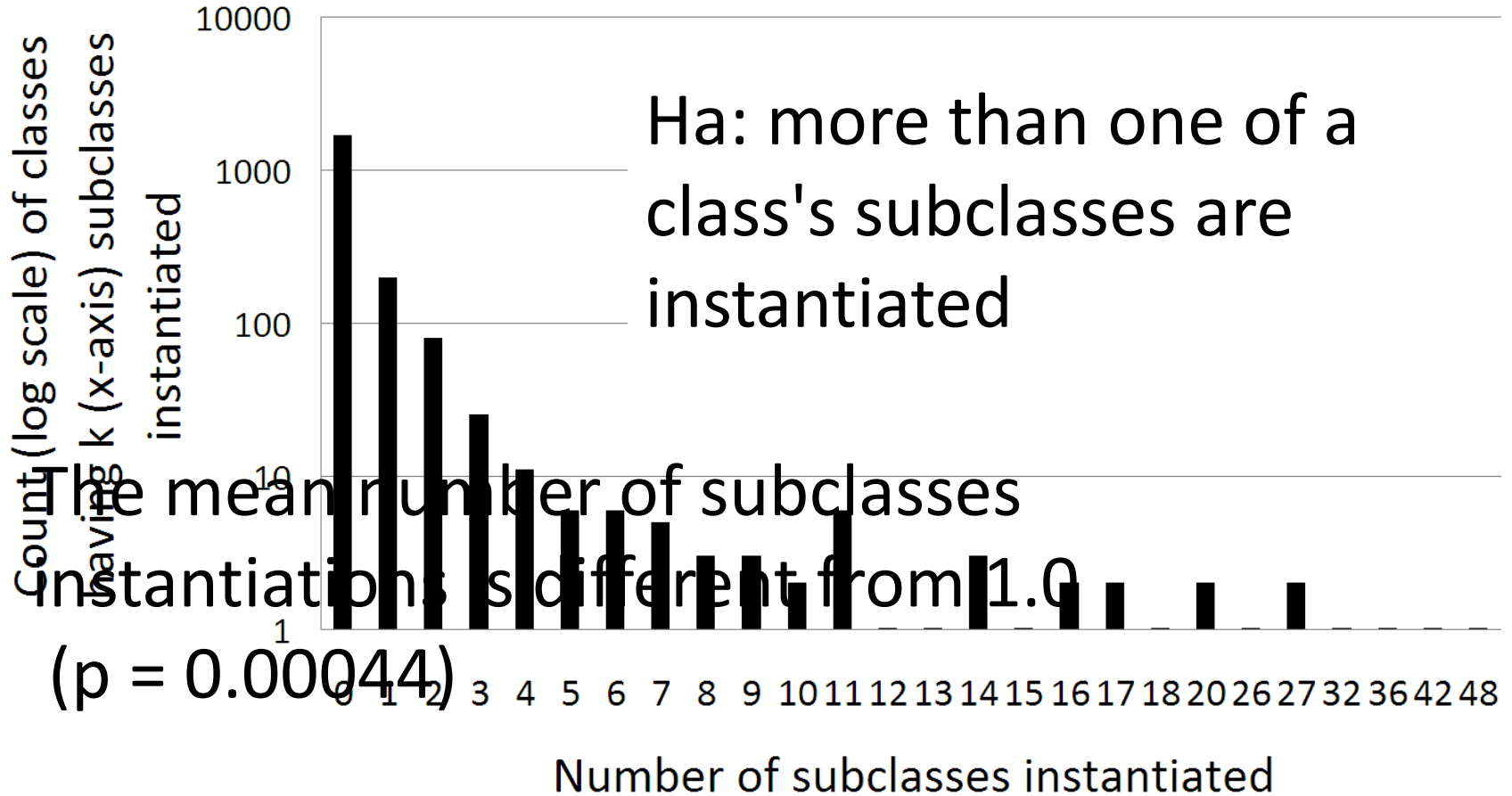   b) higher instantiation count correlate to greater skew?

Hypothesis 1

H0: only one of a class's subclass is instantiated

Ha: more than one of a class's subclasses are instantiated

Dynamic Instantiation Count

Ha: more than one of a class's subclasses are instantiated

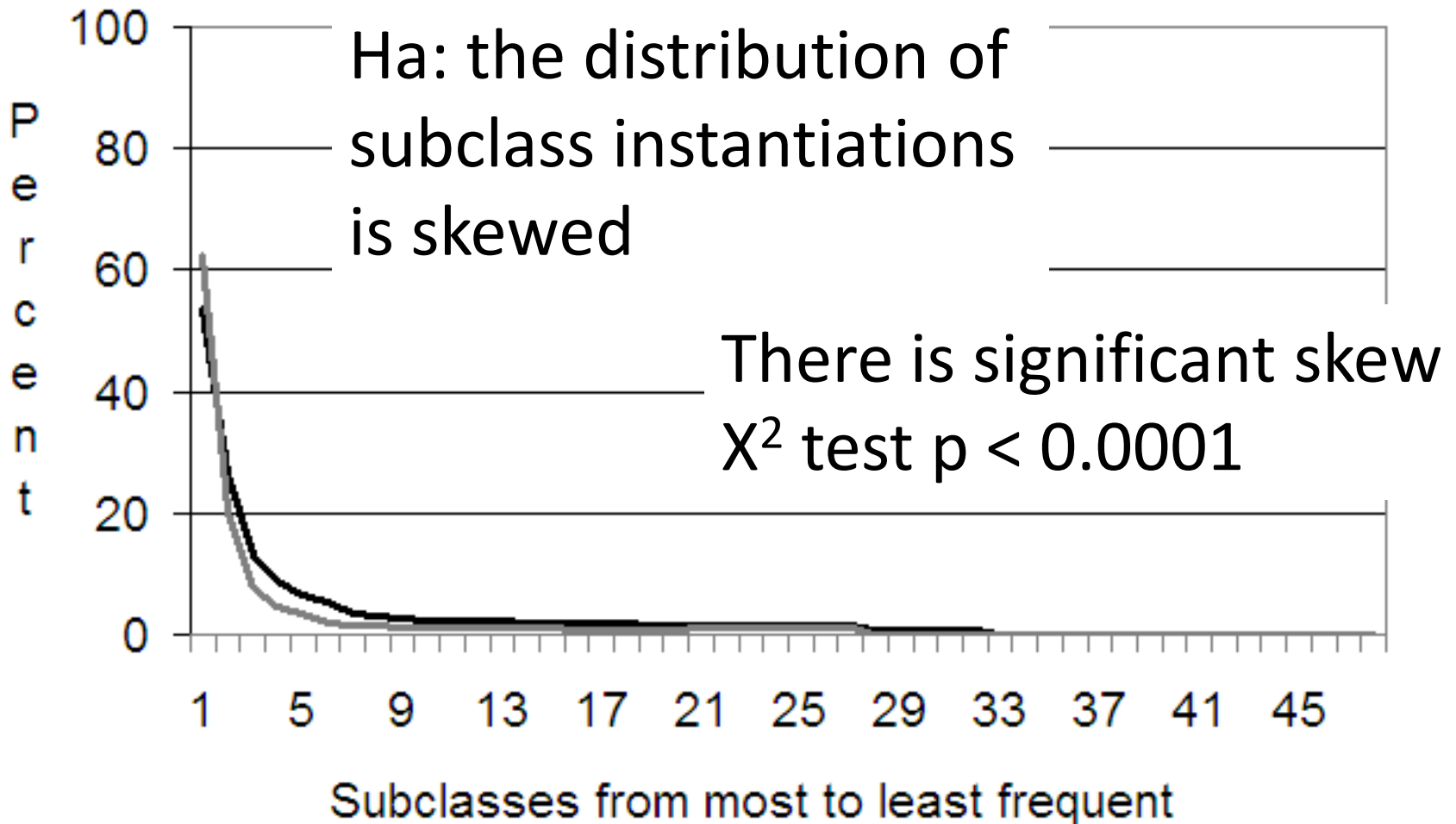The mean number of subclasses instantiations is different from 1.0 (p = 0.00044)

Hypothesis 2

H0: the distribution of subclass instantiations is uniform

Ha: the distribution of subclass instantiations is skewed
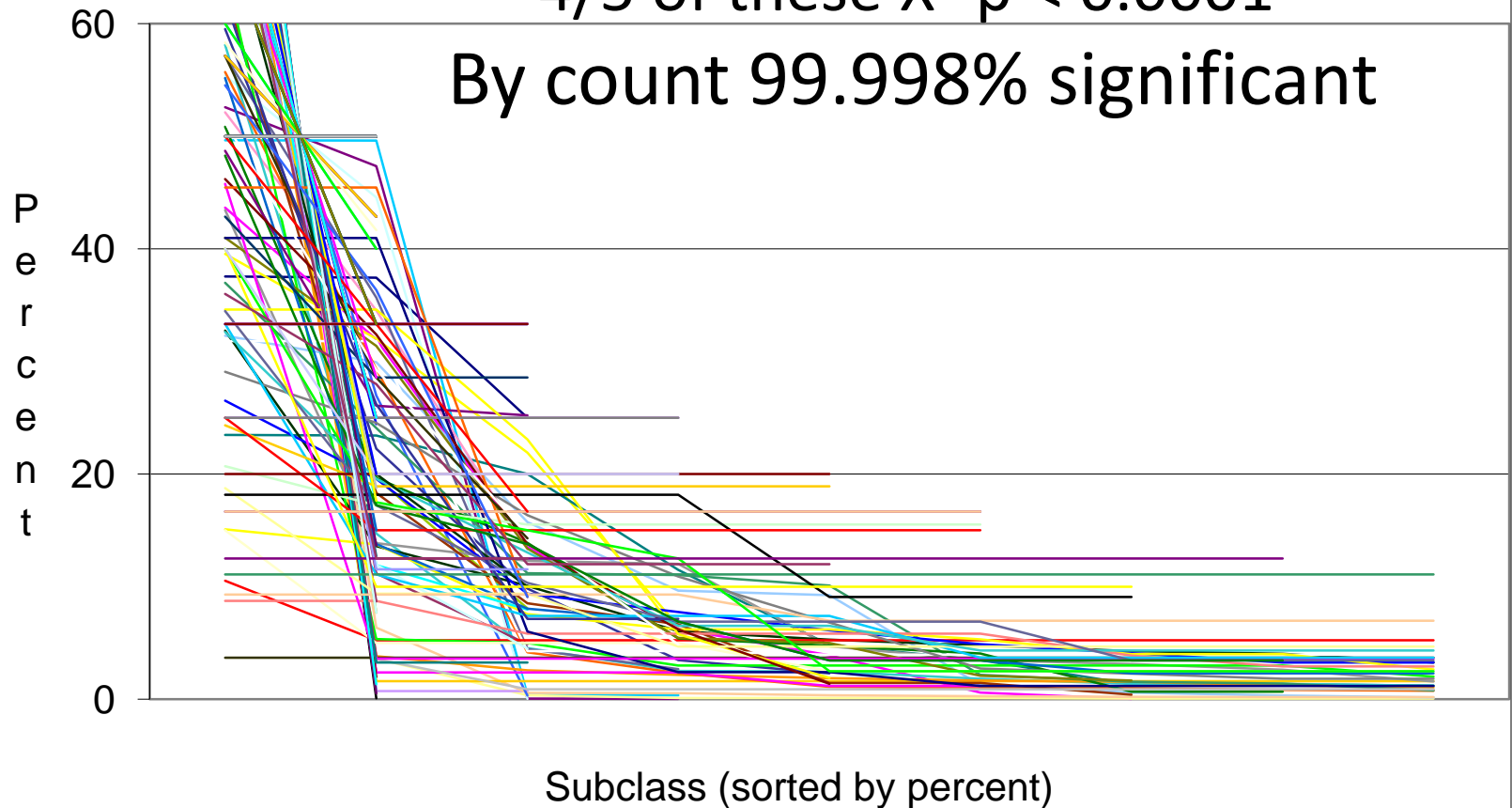
Aggregate Distribution

Ha: the distribution of subclass instantiations is skewed

There is significant skew X^2 test $p < 0.0001$

[X^2 H0 – all outcomes are equally likely]

Data per Class

By class 53% significant skew

4/5 of these X^2 p < 0.0001

By count 99.998% significant

Percent

60

40

20

0

Subclass (sorted by percent)

Hypothesis 3

Hypothesis 3a

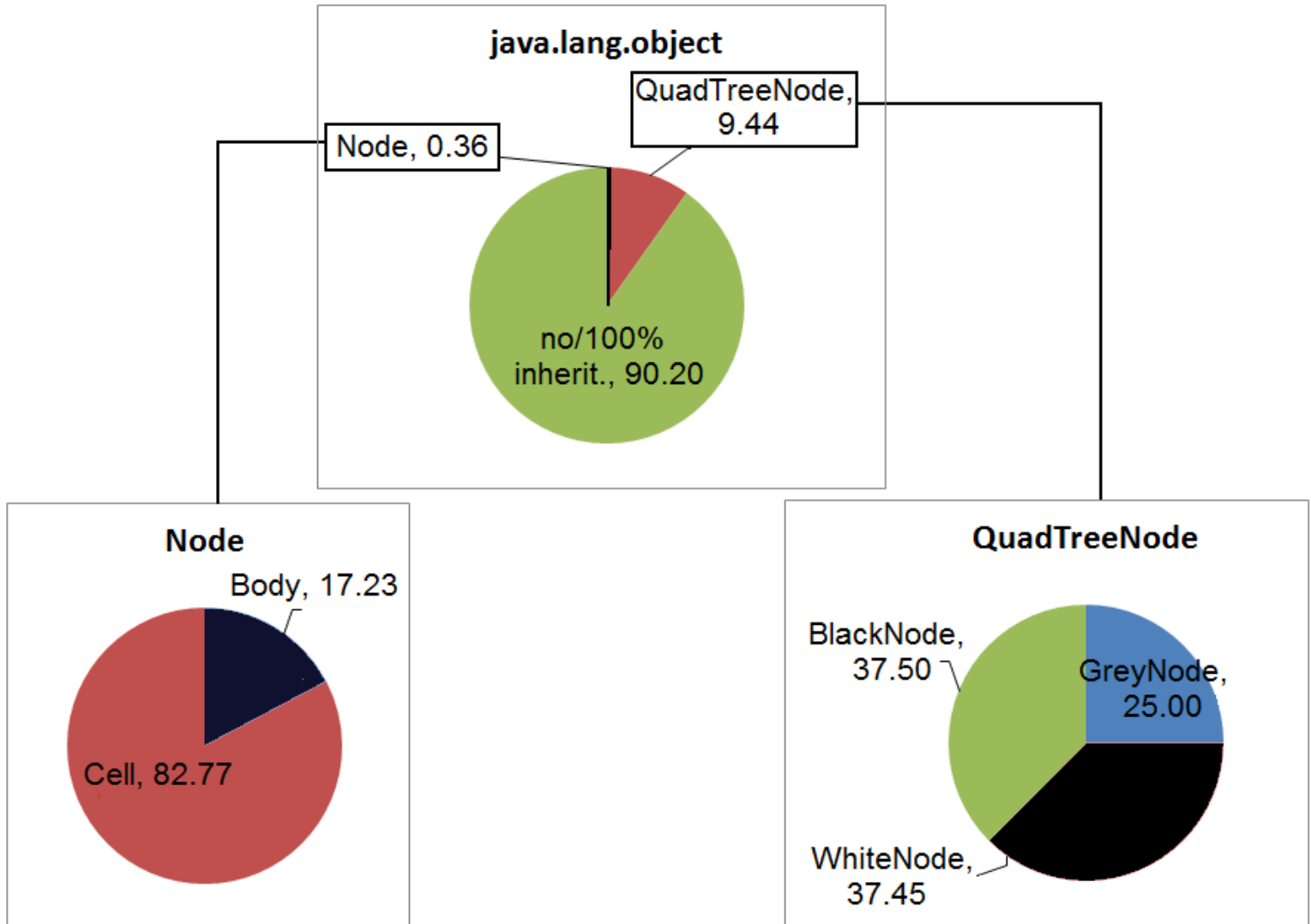Ha: one or a small number of subclasses dominate.

Hypothesis 3b

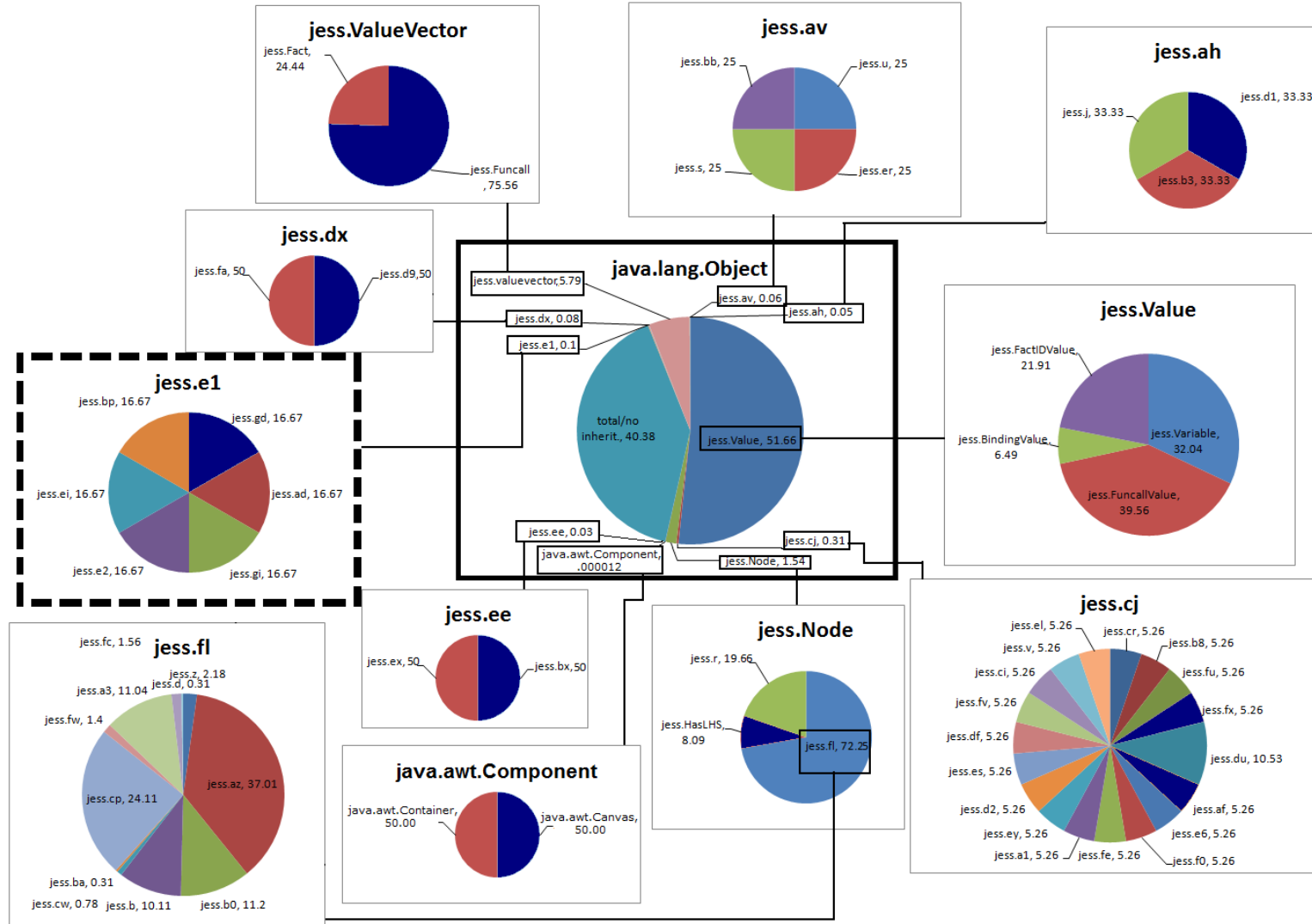Ha: higher instantiation count correlates to greater skew.

Case Studies

jolden

Jess

Thanks?Next?

Questions?

- Investigate applications
- Forecast what is typical / atypical

The Chicken came first!

- One is
  - First validate assumptions

- One is
  - First explore uses

Hypothesis 3a