The new `summary.scanone`
Karl W Broman, 27 Oct 2006

In R/qtl version 1.04, the function `summary.scanone` has been changed quite substantially. Also, the permutations with `scanone` have changed to allow the calculation of autosome- and X-chromosome-specific LOD thresholds, and to enable stratified permutation tests.

In this document, I describe the revisions and how to use the new functions. We'll first look at the `fake.f2` data as an example.

First we need to load the package and the data.

```
> library(qtl)
> data(fake.f2)
```

I'm going to use `scanone` with `method="hk"`. First I run `calc.genoprob`, and then `scanone` as before.

```
> fake.f2 <- calc.genoprob(fake.f2, step=2.5)
> out.f2 <- scanone(fake.f2, method="hk")
```

In `summary.scanone`, we can now get an indication of the number of degrees of freedom associated with the LOD scores. We use `df=TRUE`, as follows.

```
> summary(out.f2, threshold=3, df=TRUE)

Degrees of freedom: A:2  X:3

            chr  pos  lod
c1.loc27.5    1 27.5 5.12
c13.loc27.5  13 27.5 8.95
cX.loc10      X 10.0 7.20
```

There are a couple of improvements in the permutations performed by `scanone`. First, we can calculate autosome- and X-chromosome-specific LOD thresholds; this is important in this case, as the number of degrees of freedom is different for the X chromosome. Separate autosome and X chromosome permutations may be performed in `scanone` via `perm.Xsp=TRUE`. The X-chromosome-specific thresholds requires many more permutation replicates to get a threshold of equivalent precision. An increased number of permutations is chosen automatically.

Permutations can take a very long time, and so one might want to use a multi-processor computer or cluster and do multiple shorter runs in parallel. And so we have added a function `c.scanoneperm` for combining such runs together.

1

```
> operm1.f2 <- scanone(fake.f2, method="hk", n.perm=500, perm.Xsp=TRUE)
> operm2.f2 <- scanone(fake.f2, method="hk", n.perm=500, perm.Xsp=TRUE)
> operm.f2 <- c(operm1.f2, operm2.f2)
```

Getting the autosome- and X-chromosome-specific thresholds is a bit tricky, and so another improvement is the addition of the function summary.scanoneperm for calculating such thresholds. The argument alpha indicates the significance levels.

```
> summary(operm.f2, alpha=c(0.05, 0.20))


Autosome LOD thresholds (1000 permutations)
      lod
5%   3.23
20%  2.58


X chromosome LOD thresholds (37374 permutations)
      lod
5%   4.33
20%  3.42
```

Further, we may include the permutation results in the call to summary.scanone to automatically calculate thresholds and to have genome-scan-adjusted p-values displayed.

```
> summary(out.f2, perms=operm.f2, alpha=0.05, pvalues=TRUE)


  chr  pos  lod     pval
1   1 27.5 5.12 0.00103
2  13 27.5 8.95 0.00000
3   X 10.0 7.20 0.00000
```

Finally, one may prefer to do a stratified permutation test, permuting the phenotypes separately within each of the groups defined by sex and cross direction. This may be done in scanone with the argument perm.strata, which should be a numeric vector whose unique values define the separate strata.

We set of the strata as follows.

```
> sex <- fake.f2$pheno$sex
> pgm <- fake.f2$pheno$pgm
> strata <- sex + 2*pgm
> table(strata)
```

```
strata
 0  1  2  3
52 48 50 50
```

We then perform the permutation test in four pieces, and combine the results together, as follows.

```
> operm1.f2strat <- scanone(fake.f2, method="hk", n.perm=250,
+                           perm.Xsp=TRUE, perm.strata=strata)
> operm2.f2strat <- scanone(fake.f2, method="hk", n.perm=250,
+                           perm.Xsp=TRUE, perm.strata=strata)
> operm3.f2strat <- scanone(fake.f2, method="hk", n.perm=250,
+                           perm.Xsp=TRUE, perm.strata=strata)
> operm4.f2strat <- scanone(fake.f2, method="hk", n.perm=250,
+                           perm.Xsp=TRUE, perm.strata=strata)
> operm.f2strat <- c(operm1.f2strat, operm2.f2strat, operm3.f2strat,
+                    operm4.f2strat)
```

The new thresholds are as follows.

```
> summary(operm.f2strat, alpha=c(0.05, 0.20))


Autosome LOD thresholds (1000 permutations)
     lod
5%   3.42
20% 2.69


X chromosome LOD thresholds (37376 permutations)
     lod
5%   4.24
20% 3.40
```

The big changes to the `summary.scanone` function concern the case of results for multiple phenotypes. To illustrate this, we will look at the `fake.bc` data, which has two phenotypes. First we load the data.

```
> data(fake.bc)
```

Now let's run `calc.genoprob` and do a genome scan on the two phenotypes. Again, we use `method="hk"` for the sake of speed.

```
> fake.bc <- calc.genoprob(fake.bc, step=2.5)
> out.bc <- scanone(fake.bc, pheno.col=1:2, method="hk")
```

The results contain LOD scores for each of the phenotypes. By default, `summary.scanone` looks at the first of these, though it also shows the LOD score for the second phenotype at the locations of the LOD peaks for the first phenotype.

```
> summary(out.bc, threshold=3)


          chr  pos pheno1 pheno2
c2.loc32.5   2 32.5   3.52   1.87
c5.loc17.5   5 17.5   7.90   2.71
```

If we use `lodcolumn=2`, we get the analogous results, looking at the second phenotype.

```
> summary(out.bc, threshold=3, lodcolumn=2)


       chr pos pheno1 pheno2
D5M394   5 9.8   7.42    3.9
```

If we use `format="allpheno"`, we get separate rows for the peaks of each of the phenotypes.

```
> summary(out.bc, threshold=3, format="allpheno")


          chr  pos pheno1 pheno2
c2.loc32.5   2 32.5   3.52   1.87
D5M394       5  9.8   7.42   3.90
c5.loc17.5   5 17.5   7.90   2.71
```

Perhaps the most convenient output is obtained with `format="allpeaks"`, which gives a single row for each chromosome, with the maximum LOD score and its position for each of the phenotypes. A chromosome is displayed if the LOD score for at least one of the phenotypes exceeds its threshold. The `threshold` argument can be a single threshold, applied to all phenotypes, or we can give a vector with separate thresholds for each of the LOD score columns.

```
> summary(out.bc, threshold=c(3,2.5), format="allpeaks")


  chr  pos pheno1  pos pheno2
1   2 32.5   3.52 37.5   1.91
2   5 17.5   7.90  9.8   3.90
```

A permutation test may be performed as before. Since the `fake.bc` data has only autosomal data, use of `perm.Xsp=TRUE` would be ignored.

```
> operm.bc <- scanone(out.bc, pheno.col=1:2, method="hk", n.perm=1000)
```

We can again use `summary` to get LOD thresholds

```
> summary(operm.bc, alpha=0.05)


LOD thresholds (1000 permutations)
   pheno1 pheno2
5%   2.57   2.55
```

And again these can be used in `summary.scanone` to calculate thresholds and get genome-scan-adjusted p-values.

```
> summary(out.bc, perms=operm.bc, alpha=0.05, format="allpeaks",
+         pvalues=TRUE)


  chr  pos pheno1  pval  pos pheno2  pval
1   2 32.5   3.52 0.005 37.5   1.91 0.195
2   5 17.5   7.90 0.000  9.8   3.90 0.002
```