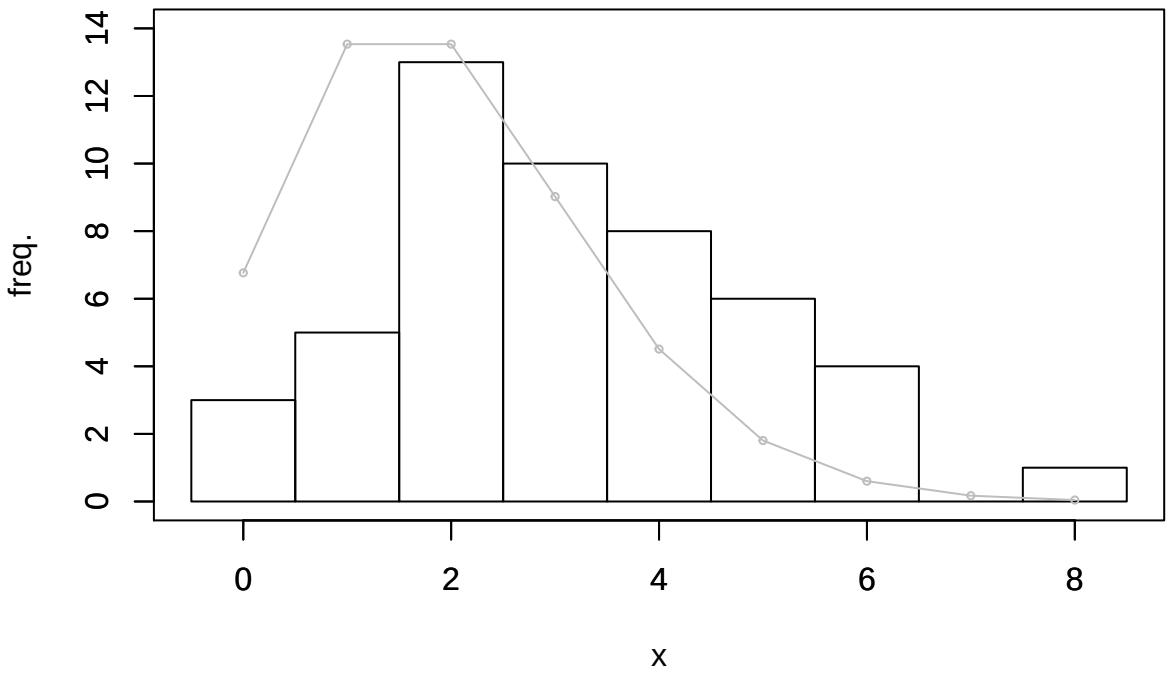
note11

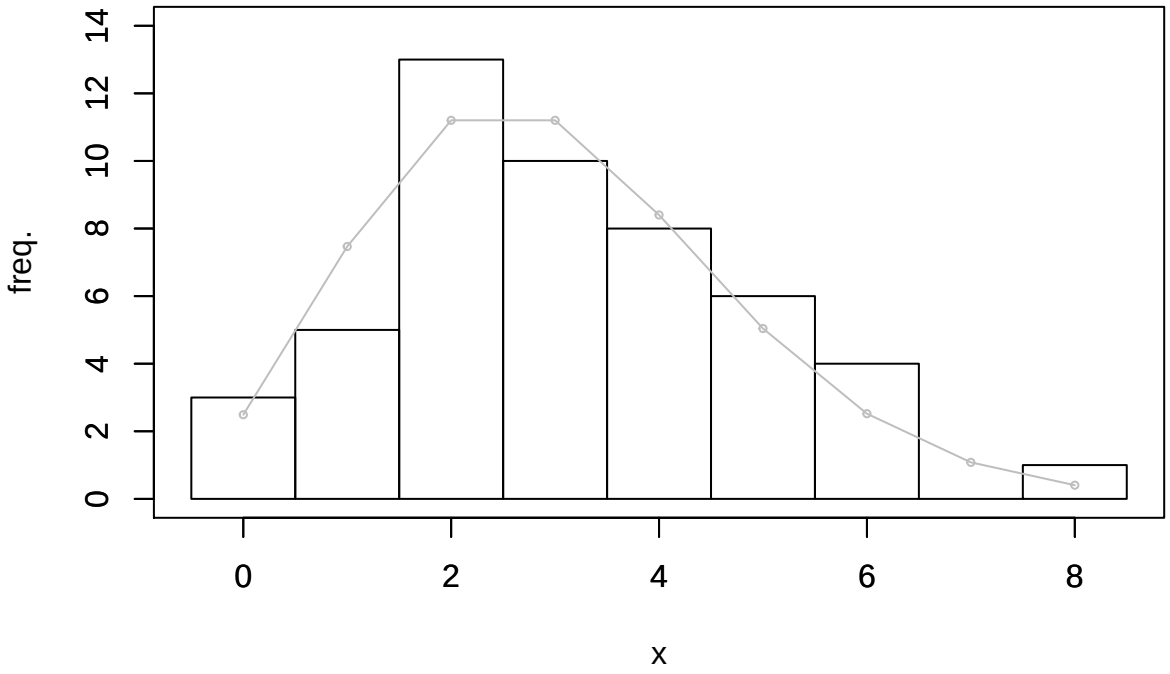
Matts966

2018/01/30

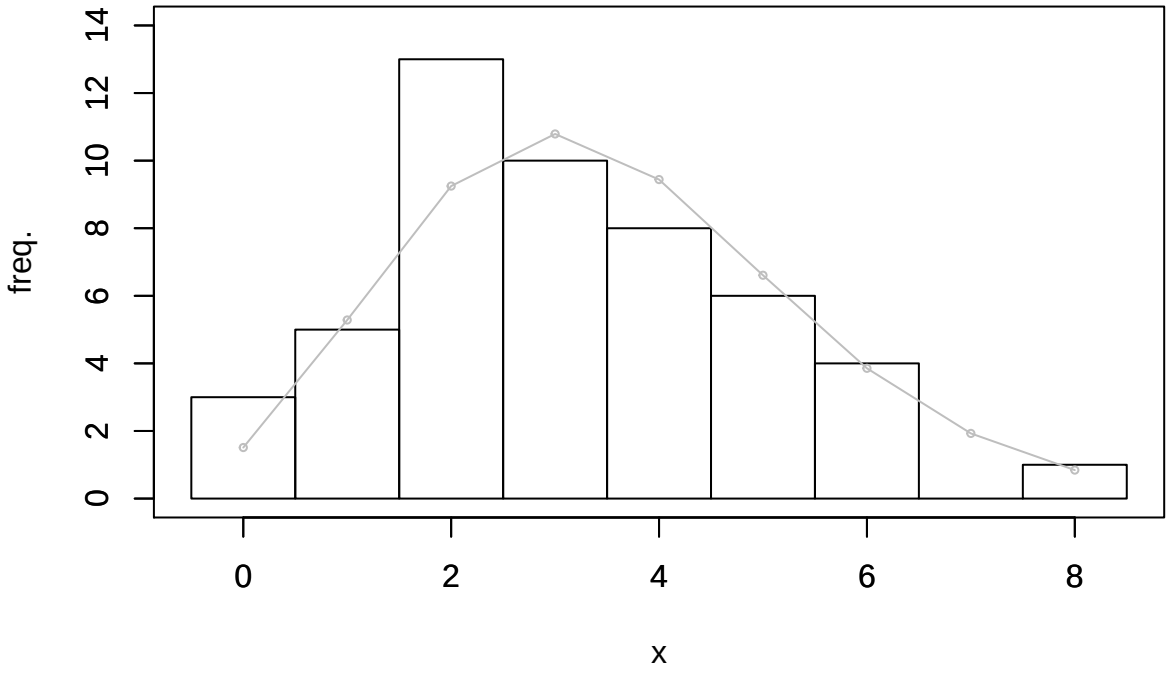
```r
# seed をセットして再現性を持たせる
set.seed(123)

# type=o はオーバーラップ
l <- 3
d <- rpois(50, lambda=l)
hist(d, ylim=c(0, 14), breaks=c(-0.5:8.5), xlab="x", ylab="freq.", main="")
par(new=T)
x <- seq(0, 8, 1)
y <- 50*dpois(x, lambda=2)
plot(x, y, xlab="", ylab="", xlim=c(-0.5, 8.5), ylim=c(0, 14), type="o", col="grey", cex=0.5)
```



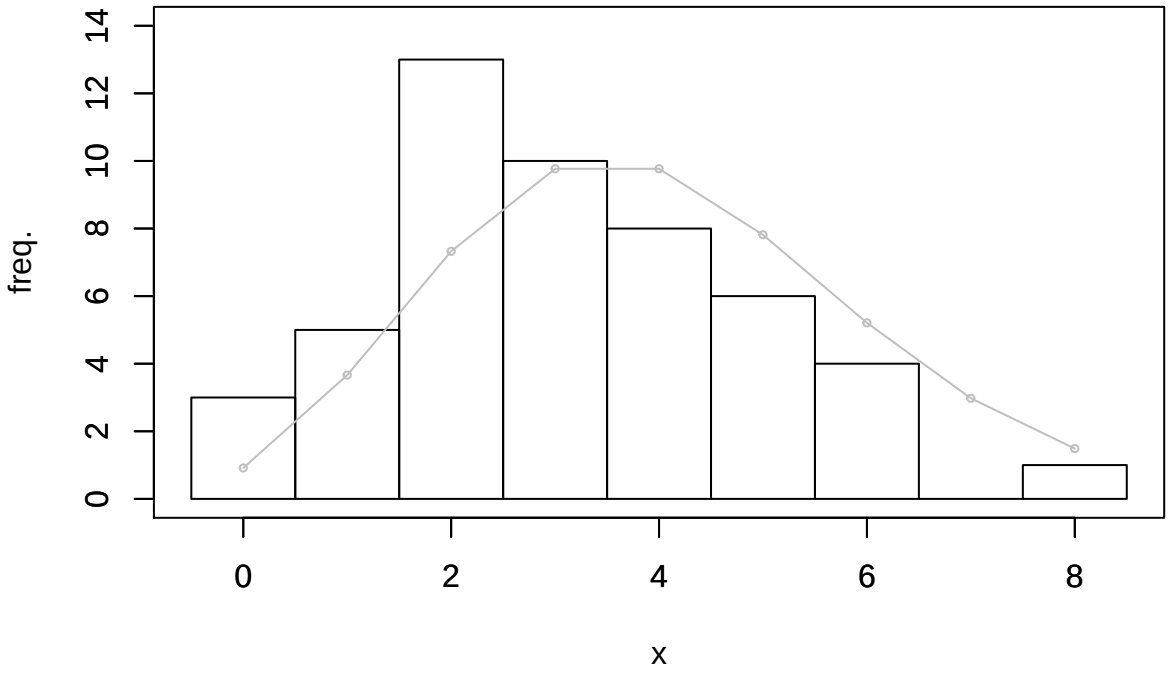

```r
hist(d, ylim=c(0, 14), breaks=c(-0.5:8.5), xlab="x", ylab="freq.", main="")
par(new=T)
x <- seq(0, 8, 1)
y <- 50*dpois(x, lambda=3)
plot(x, y, xlab="", ylab="", xlim=c(-0.5, 8.5), ylim=c(0, 14), type="o", col="grey", cex=0.5)
```




```
hist(d, ylim=c(0, 14), breaks=c(-0.5:8.5), xlab="x", ylab="freq.", main="")
par(new=T)
x <- seq(0, 8, 1)
y <- 50*dpois(x, lambda=3.5)
plot(x, y, xlab="", ylab="", xlim=c(-0.5, 8.5), ylim=c(0, 14), type="o", col="grey", cex=0.5)
```




```
hist(d, ylim=c(0, 14), breaks=c(-0.5:8.5), xlab="x", ylab="freq.", main="")
par(new=T)
x <- seq(0, 8, 1)
y <- 50*dpois(x, lambda=4)
plot(x, y, xlab="", ylab="", xlim=c(-0.5,8.5), ylim=c(0, 14), type="o", col="grey", cex=0.5)
```




```
set.seed(123)
l <- 3
d <- rpois(50, lambda=l)

logL <- sum(dpois(d, 2.0, log=T))
logL
```

```
## [1] -110.5617
```

```
logL <- sum(dpois(d, 3.0, log=T))
logL
```

```
## [1] -97.71459
```

```
logL <- sum(dpois(d, 3.5, log=T))
logL
```

```
## [1] -98.82124
```

```
logL <- sum(dpois(d, 4.0, log=T))
logL
```

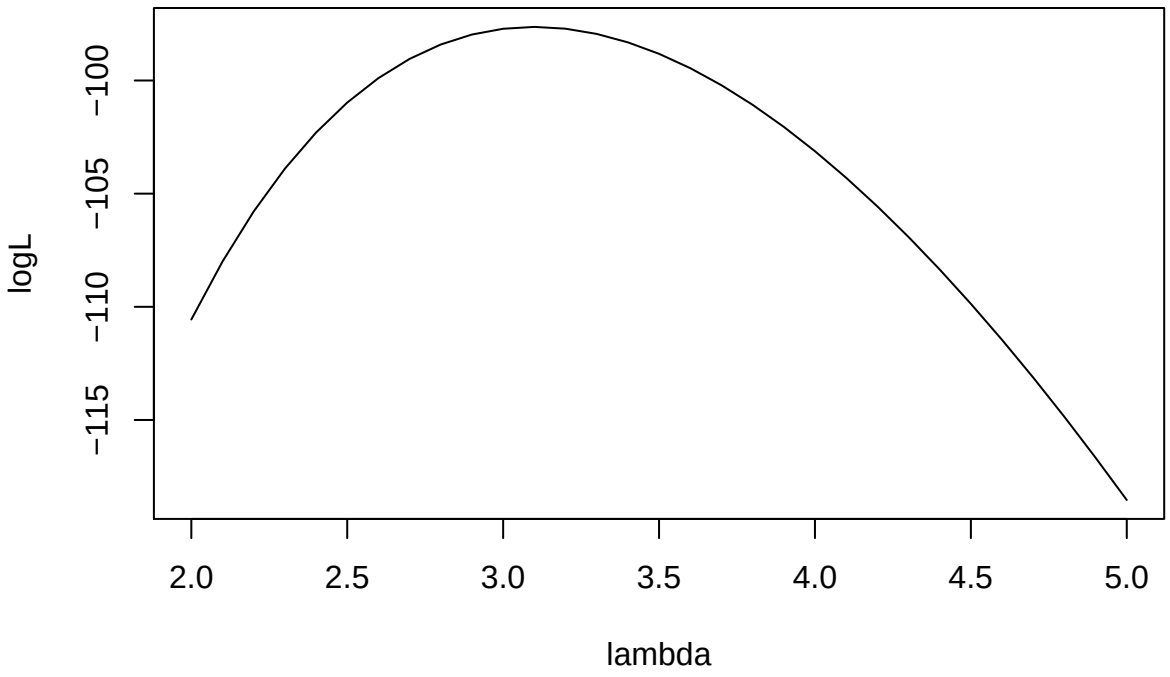
```
## [1] -103.1239
```

```
# AICを計算して最小を提示
AIC = -2 * (-97.71459) + 2 * 2
AIC
```

```
## [1] 199.4292
```

```
# よって lambda3.0 で AIC最小
```

```
# figure 10.7
set.seed(123)
l <- 3
d <- rpois(50, lambda=l)

logL <- function(m) sum(dpois(d, m, log=T))
lambda <- seq(2, 5, 0.1)
plot(lambda, sapply(lambda, logL), type="l", xlim=c(2, 5), xlab="lambda", ylab="logL")
```




```
setwd("/Users/masahiromatsui/Dropbox/R/R で学ぶ統計学入門図版作成用（改訂版）/付録/")
d <- read.csv("table10-2.csv")
result <- glm(cbind(d$dead, 1-d$dead) ~ d$dose, family=binomial(logit))
logLik(result)
```

```
## 'log Lik.' -6.781799 (df=2)
```

```
AIC = -2 * (-6.781799) + 2 * 2
AIC
```

```
## [1] 17.5636
```

```
setwd("/Users/masahiromatsui/Dropbox/R/R で学ぶ統計学入門図版作成用（改訂版）/付録/")
d <- read.csv("table10-3.csv")
result <- glm(d$flw ~ d$wt, family=poisson)
logLik(result)
```

```
## 'log Lik.' -82.84586 (df=2)
```

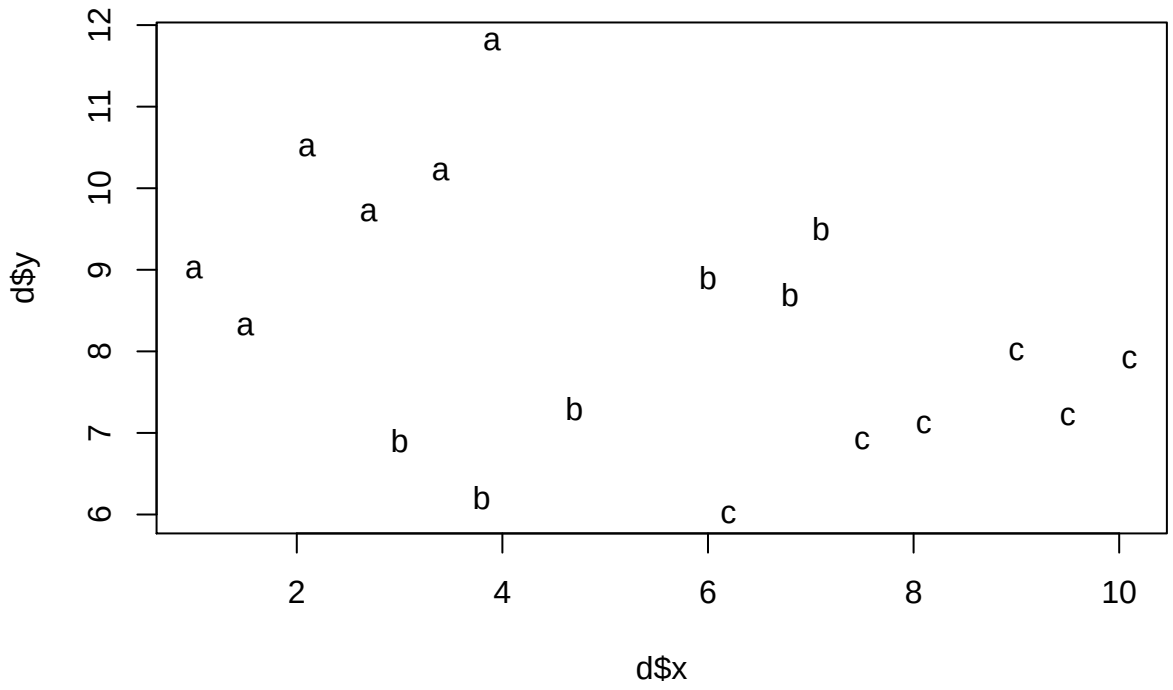
```
AIC = -2 * (-82.84586) + 2 * 2
AIC
```

```
## [1] 169.6917
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
setwd("/Users/masahiromatsui/Dropbox/R/R で学ぶ統計学入門図版作成用（改訂版）/付録/")

d <- read.csv("table11-1.csv")
plot(d$y ~ d$x, pch=as.character(d$block))
```




```
res.1 <- glmer(d$y ~d$x + (1|d$block), family=gaussian(link = identity))
```

```
## Warning in glmer(d$y ~ d$x + (1 | d$block), family = gaussian(link =
## identity)): calling glmer() with family=gaussian (identity link) as a
## shortcut to lmer() is deprecated; please call lmer() directly
```

```
res.2 <- lmer(d$y ~ d$x + (1|d$block))
```

```
summary(res.1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: d$y ~ d$x + (1 | d$block)
##
## REML criterion at convergence: 45.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.6493 -0.5949  0.1023  0.6671  1.5821
##
## Random effects:
```

```
##  Groups   Name        Variance Std.Dev.
##  d$block  (Intercept) 11.0023  3.3170
##  Residual              0.3631  0.6026
## Number of obs: 18, groups:  d$block, 3
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   4.8738     2.0063   2.429
## d$x           0.6470     0.1085   5.962
##
## Correlation of Fixed Effects:
##     (Intr)
## d$x -0.290
```

```
summary(res.2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: d$y ~ d$x + (1 | d$block)
##
## REML criterion at convergence: 45.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.6493 -0.5949  0.1023  0.6671  1.5821
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  d$block  (Intercept) 11.0023  3.3170
##  Residual              0.3631  0.6026
## Number of obs: 18, groups:  d$block, 3
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   4.8738     2.0063   2.429
## d$x           0.6470     0.1085   5.962
##
## Correlation of Fixed Effects:
##     (Intr)
## d$x -0.290
```

【考察】AIC は -2 * 最大対数尤度 + 2 * 最尤推定した自由パラメータの個数で表され、加点対象、ペナルティをを演算していることはわかったのだが、なぜ最尤推定した自由パラメータの個数 - 最大対数尤度として大きい程よい値に設定しないのか、そもそも単純な足し引きで指標を決定して良いのか、理解が至らなかったので、これから勉強したい。