How good is the Chebychev bound ?

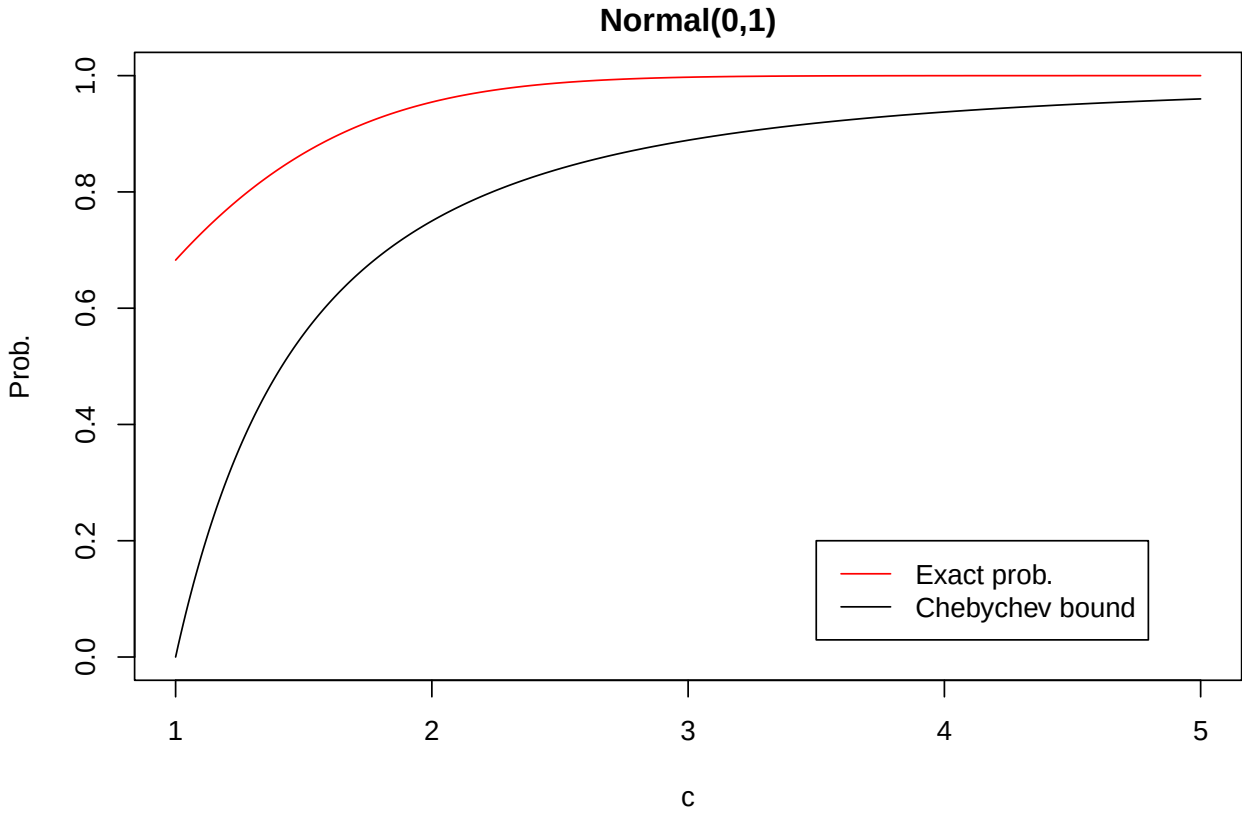
Recall the following probability inequalities:

- **Markov's inequality:** Suppose that $g(Y)$ is a function of random variable Y such that $\mathbb{E}[|g(Y)|] < \infty$. For any $k > 0$, it can be shown that
$$P(|g(Y)| \geq k) \leq \frac{\mathbb{E}[|g(Y)|]}{k}.$$
or equivalently
$$P(|g(Y)| < k) \geq 1 - \frac{\mathbb{E}[|g(Y)|]}{k}.$$
This is **Markov's inequality**.

- **Chebychev's inequality:** In Markov's inequality, consider the function $g(t) = (t-\mu)^2$. We have for any $k > 0$ that
$$P(|(Y-\mu)^2| < k) \geq 1 - \frac{\mathbb{E}[|(Y-\mu)^2|]}{k}.$$
which can be more simply written as
$$P((Y-\mu)^2 < k) \geq 1 - \frac{\mathbb{E}[(Y-\mu)^2]}{k}.$$
But $\mathbb{E}[(Y-\mu)^2] = \sigma^2$, and if we choose to write $k = c^2\sigma^2$, the inequality becomes
$$P((Y-\mu)^2 < c^2\sigma^2) \geq 1 - \frac{\mathbb{E}[(Y-\mu)^2]}{c^2\sigma^2} = 1 - \frac{1}{c^2}$$
The left hand side of the inequality is equal to
$$P(\sqrt{(Y-\mu)^2} < \sqrt{c^2\sigma^2}) = P(|Y-\mu| < c\sigma)$$
so the inequality becomes
$$P(|Y-\mu| < c\sigma) \geq 1 - \frac{1}{c^2}.$$
or equivalently
$$P(|Y-\mu| \geq c\sigma) \leq \frac{1}{c^2}.$$
This is **Chebychev's inequality**: for **any** distribution with finite expectation μ and variance σ^2, we have that
$$P(\mu - c\sigma < Y < \mu + c\sigma) \geq 1 - \frac{1}{c^2}$$
which bounds how much probability can be present in the tails of the distribution. The bound is redundant if $c \leq 1$.

1. **Normal:** in the case of the Normal distribution studying the standard case is sufficient. For Chebychev's inequality, we assess the bound
$$P(-c < Y < c) \geq 1 - \frac{1}{c^2}$$

```
mu<-0
sig<-1
cval<-seq(1,5,by=0.01)
pval<-pnorm(cval)-pnorm(-cval)
par(mar=c(4,4,2,0))
plot(cval,1-1/cval^2,type='l',xlab='c',ylab='Prob.',ylim=range(0,1))
lines(cval,pval,col='red')
legend(3.5,0.2,c('Exact prob.','Chebychev bound'),col=c('red','black'),lty=1)
title('Normal(0,1)')
```



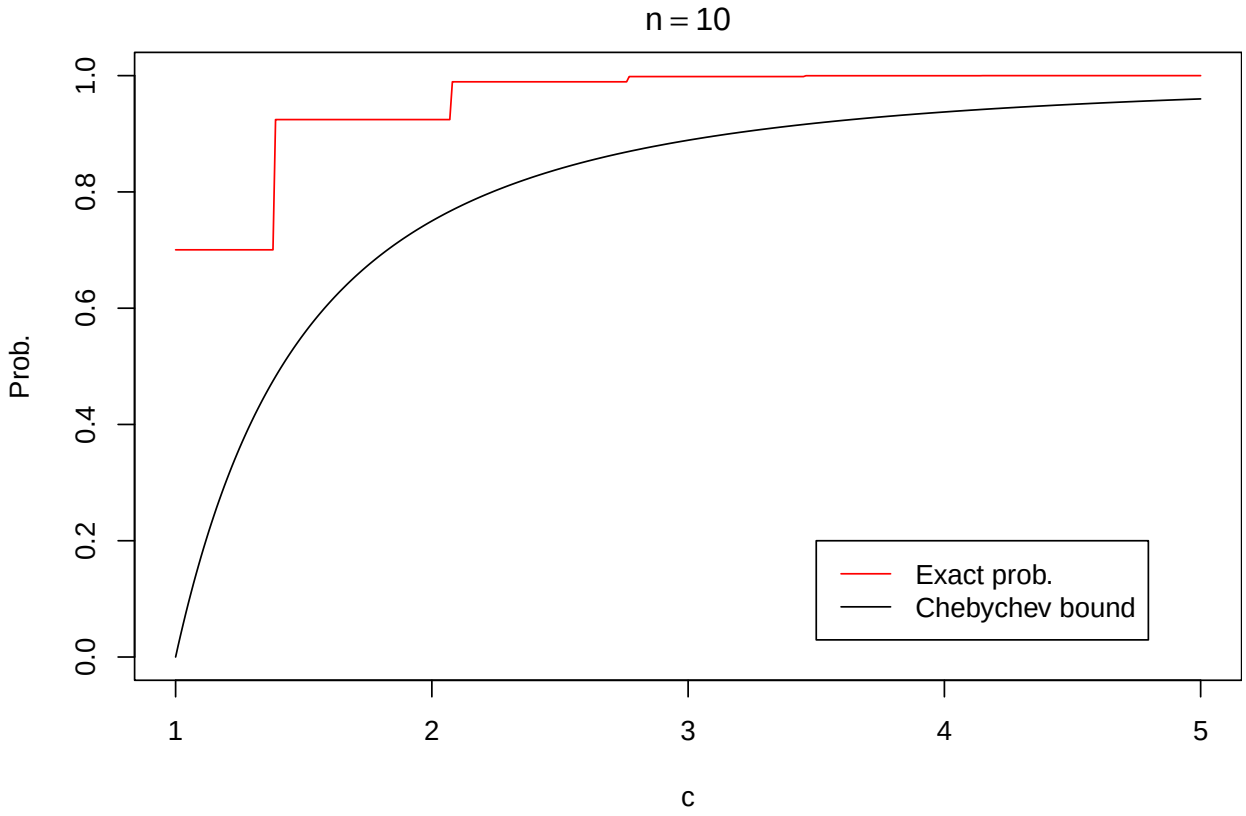
2. **Binomial:** For the $Binomial(n,p)$, we assess the bound

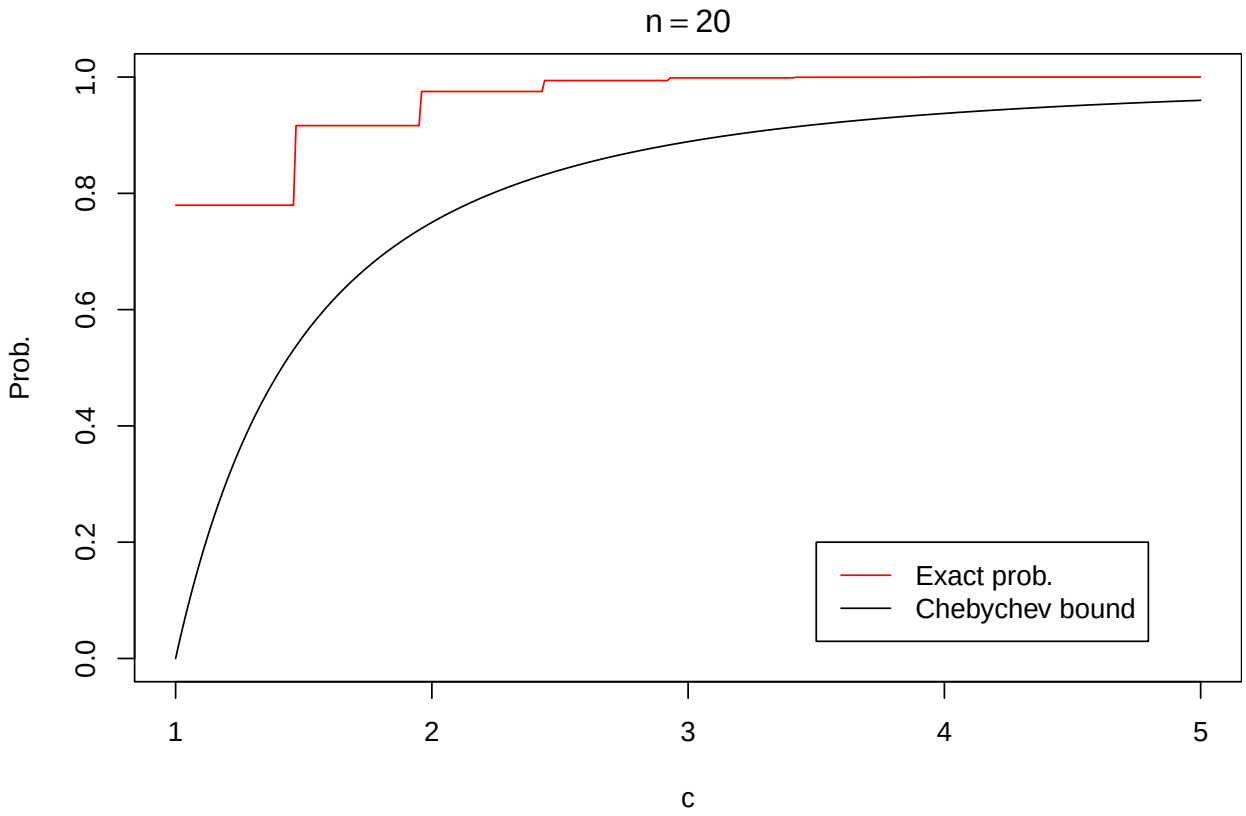
$$P(np - c\sqrt{np(1-p)} < Y < np + c\sqrt{np(1-p)}) \geq 1 - \frac{1}{c^2}$$

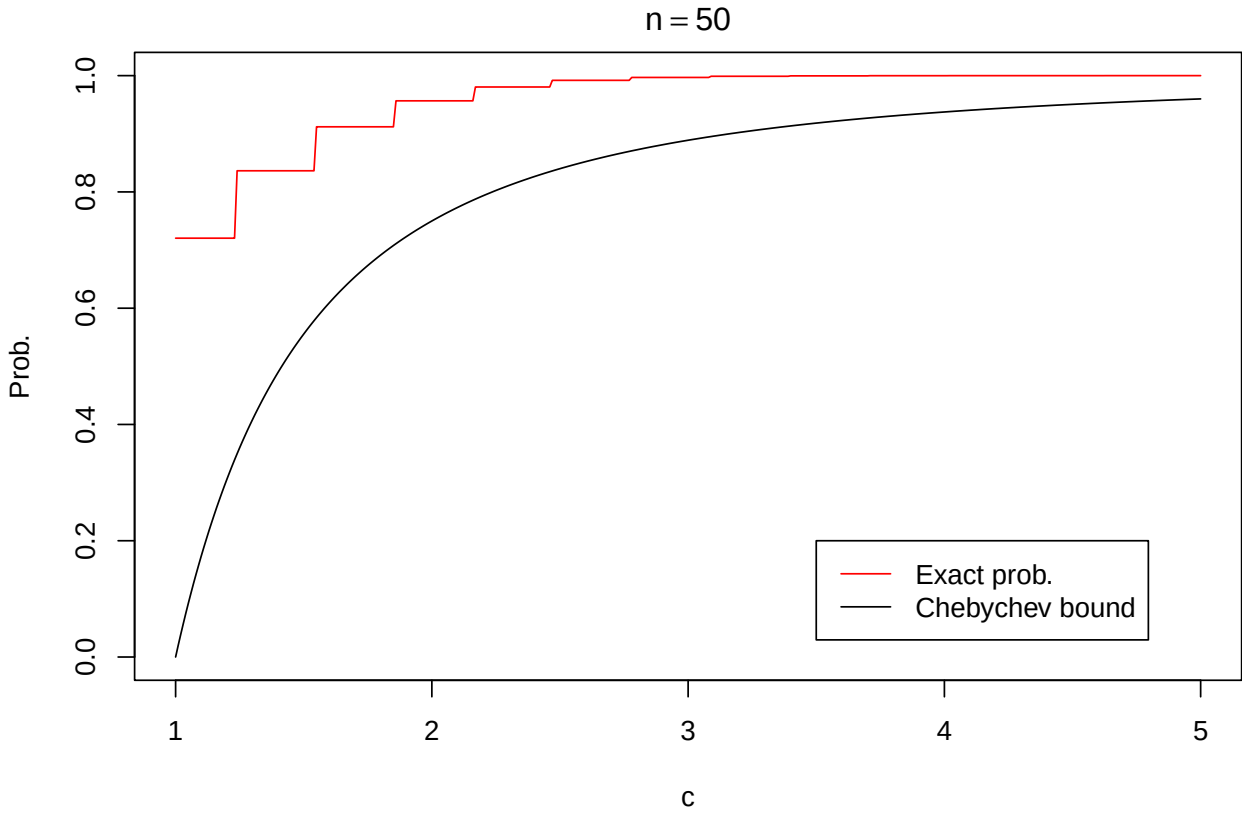
or equivalently

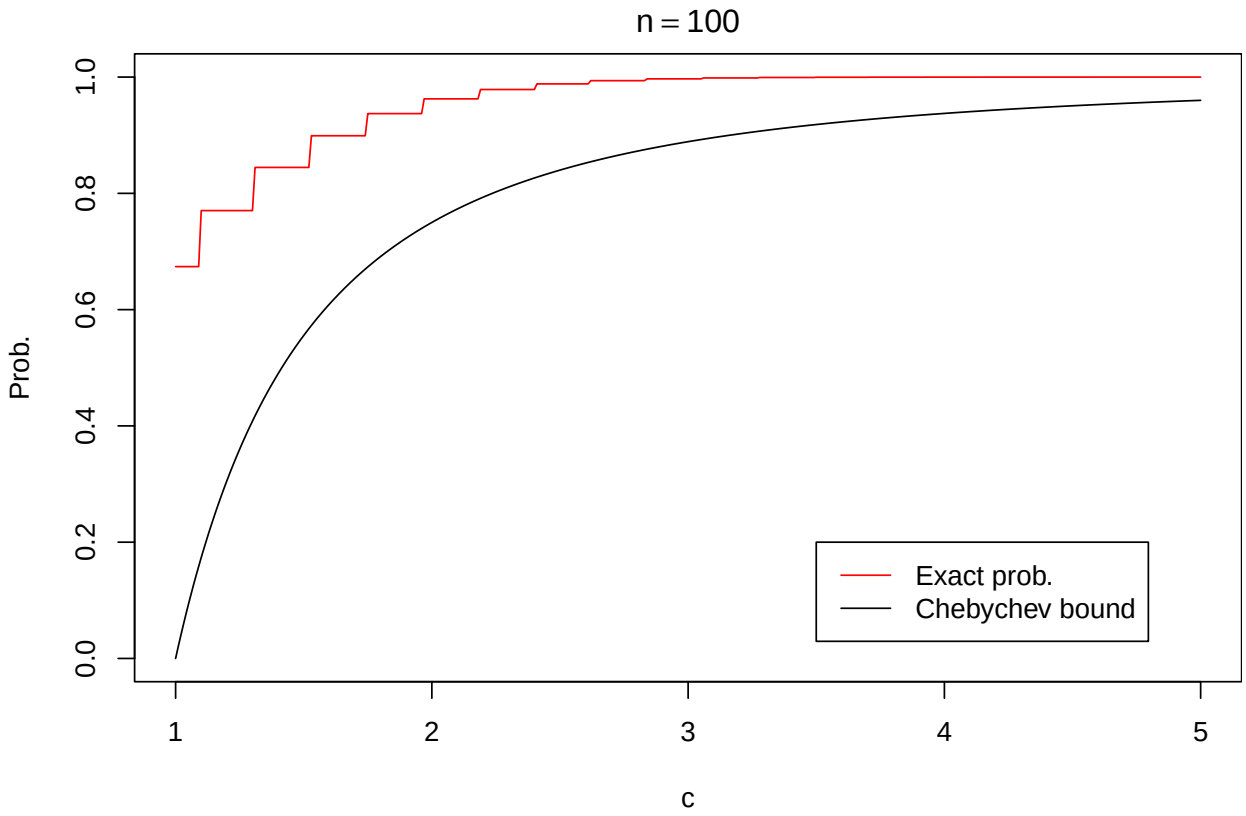
$$P\left(-c < \frac{Y - np}{\sqrt{np(1-p)}} < c\right) \geq 1 - \frac{1}{c^2}$$

```
p<-0.3
nvec<-c(10,20,50,100)
muvec<-nvec*p
sigvec<-sqrt(nvec*p*(1-p))
cval<-seq(1,5,by=0.01)
par(mar=c(4,4,2,0))
for(i in 1:4){
    pval<-pbinom(muvec[i]+cval*sigvec[i],nvec[i],p)-pbinom(muvec[i]-cval*sigvec[i],nvec[i],p)
    plot(cval,1-1/cval^2,type='l',xlab='c',ylab='Prob.',ylim=range(0,1))
    lines(cval,pval,col='red')
    legend(3.5,0.2,c('Exact prob.','Chebychev bound'),col=c('red','black'),lty=1)
    title(substitute(n==nv,list(nv=nvec[i])))
}
```


n = 10
Prob.
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
c
Exact prob.
Chebychev bound


n = 20
Prob.
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
c
Exact prob.
Chebychev bound


n = 50
Prob.
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
c
Exact prob.
Chebychev bound


n = 100
Prob.
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
c
Exact prob.
Chebychev bound

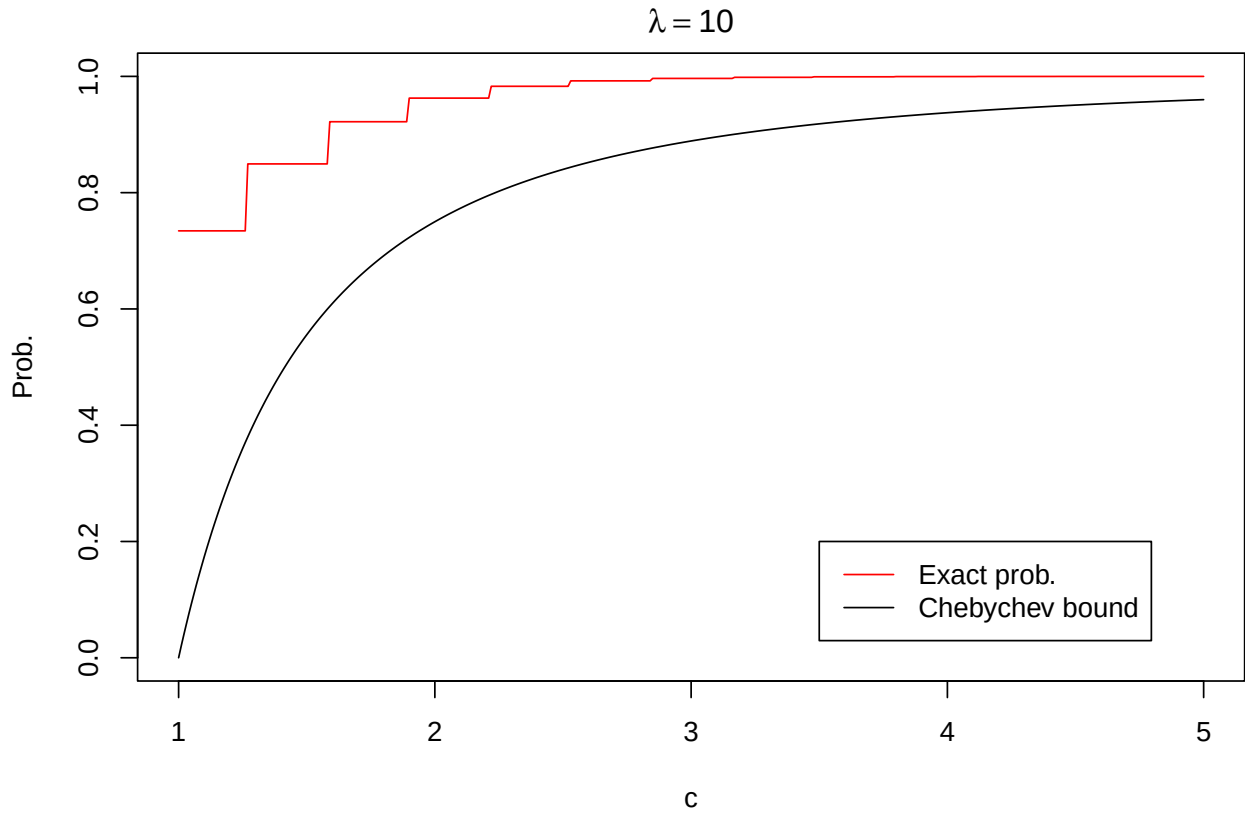
3. **Poisson:** For the $Poisson(\lambda)$, we assess the bound

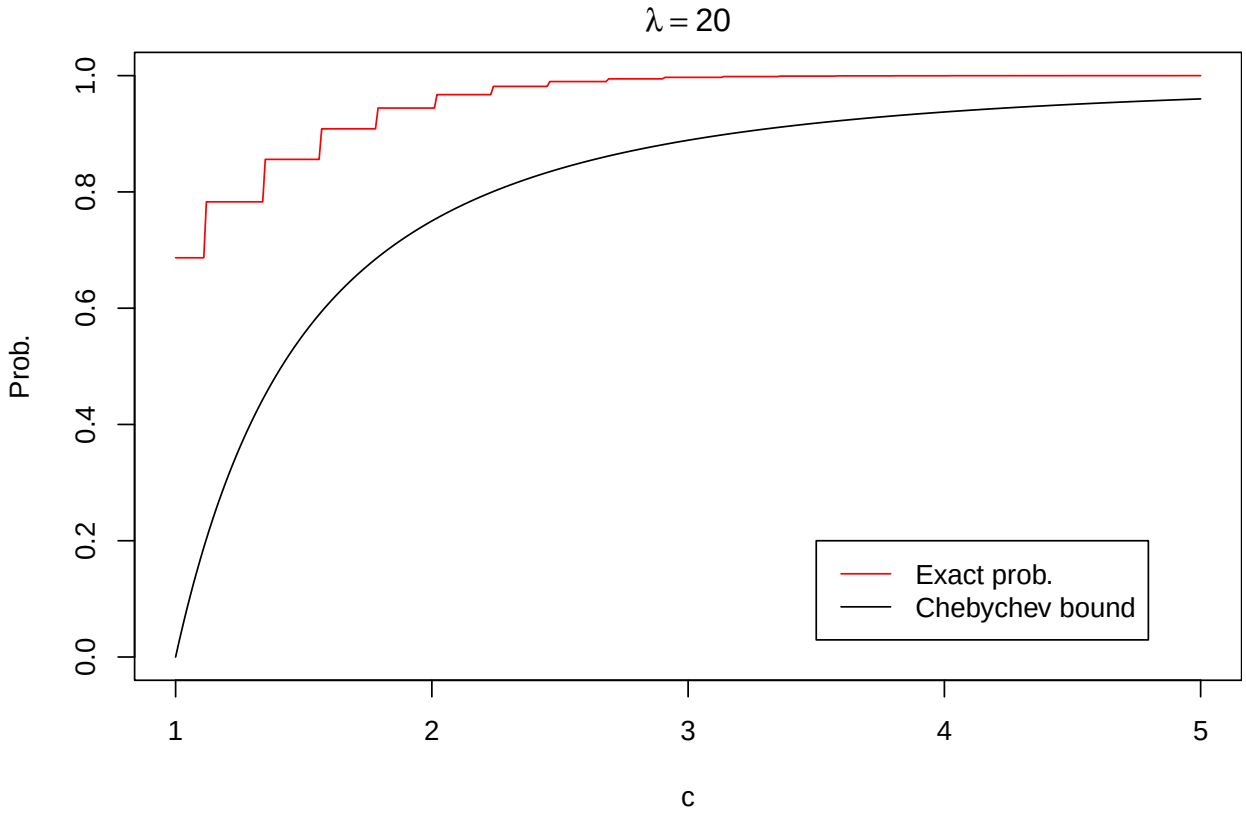
$$P(\lambda - c\sqrt{\lambda} < Y < \lambda + c\sqrt{\lambda}) \geq 1 - \frac{1}{c^2}$$

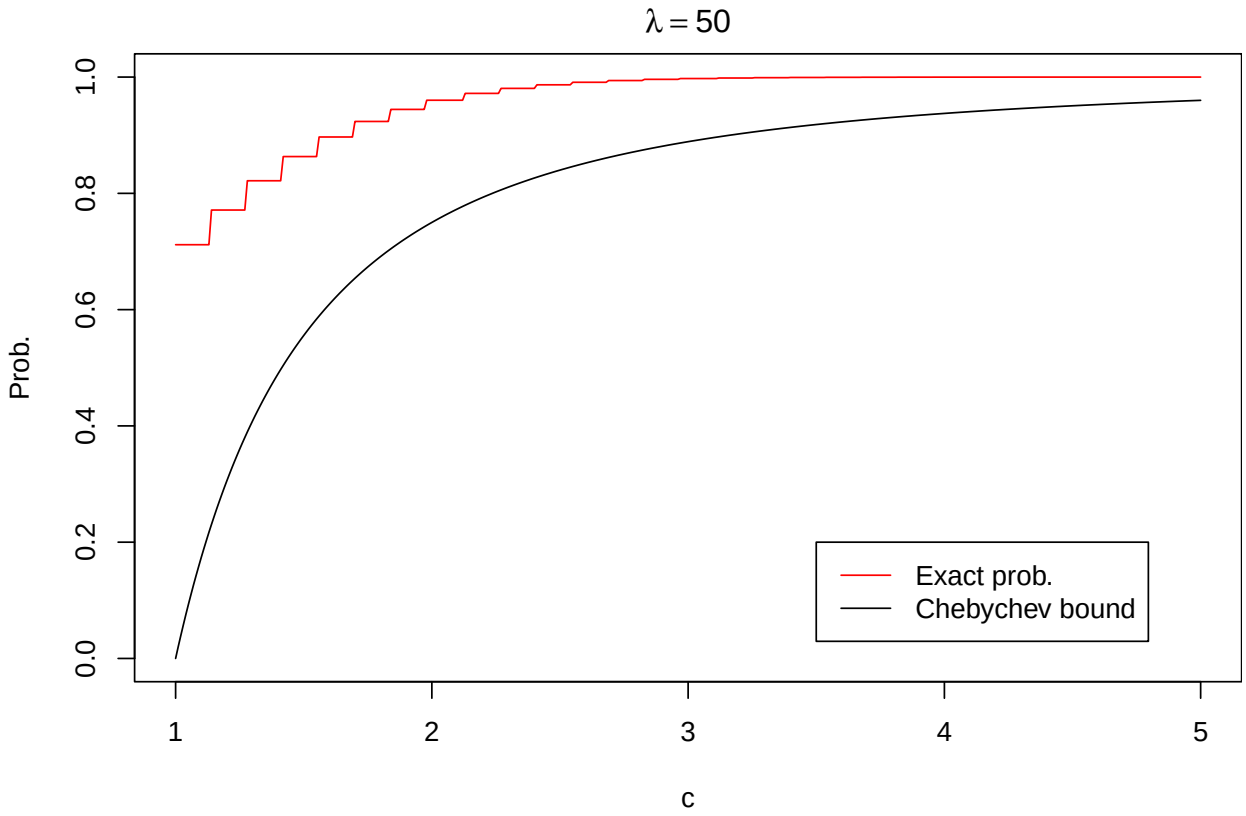
or equivalently

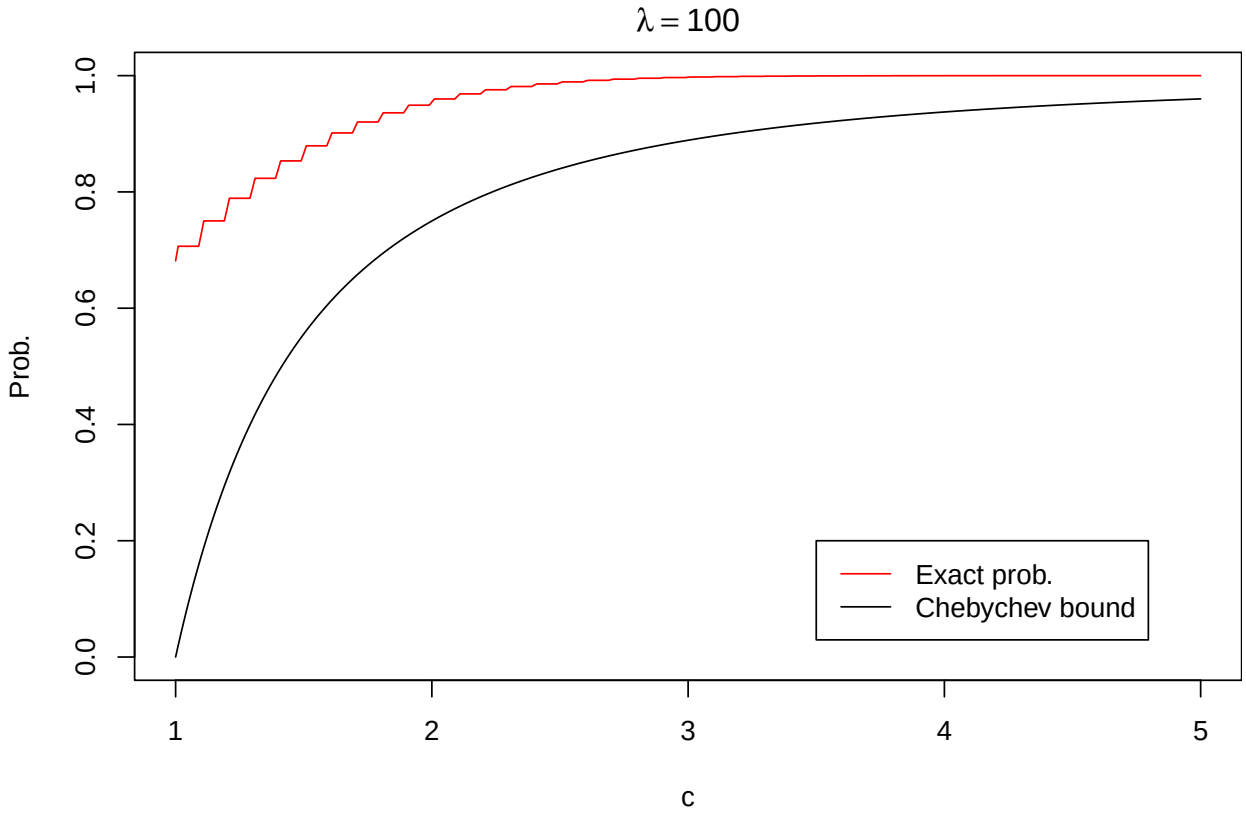
$$P\left(-c < \frac{Y - \lambda}{\sqrt{\lambda}} < c\right) \geq 1 - \frac{1}{c^2}$$

```
p<-0.3
lamvec<-c(10,20,50,100)
muvec<-lamvec
sigvec<-sqrt(lamvec)
cval<-seq(1,5,by=0.01)
par(mar=c(4,4,2,0))
for(i in 1:4){
    pval<-ppois(muvec[i]+cval*sigvec[i],lamvec[i])-ppois(muvec[i]-cval*sigvec[i],lamvec[i])
    plot(cval,1-1/cval^2,type='l',xlab='c',ylab='Prob.',ylim=range(0,1))
    lines(cval,pval,col='red')
    legend(3.5,0.2,c('Exact prob.','Chebychev bound'),col=c('red','black'),lty=1)
    title(substitute(lambda==lv,list(lv=lamvec[i])))
}
```




λ = 20
Prob.
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
c
Exact prob.
Chebychev bound


λ = 50
Prob.
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
c
Exact prob.
Chebychev bound



In all of these cases, the bound is not very tight, although it improves as c gets larger.