刘程华 2018011687 计91

3.1 Verify the expressions given for the mean and variance in (3.2) for statistics based on sums S_{+} or S_{-} of signed scores $s_i, 1=1,2,\ldots,n$ where the sign associated with each s_i is equally likely to be plus or minus.

Any s_i is included in S_{+} if it has a plus sign, otherwise it makes zero contribution. Probability it has plus sign is 0.5. Thus $\mathrm{E}\left(s_i\right)=s_i/2$ and $\operatorname{Var}(s_i)=s_i^2/4$ since the s_i are independent it follows that

$\mathrm{E}\left(S_{+}\right)=\sum_{i=1}^{n}\left(s_i/2\right)=\frac{n(n+1)}{4}$

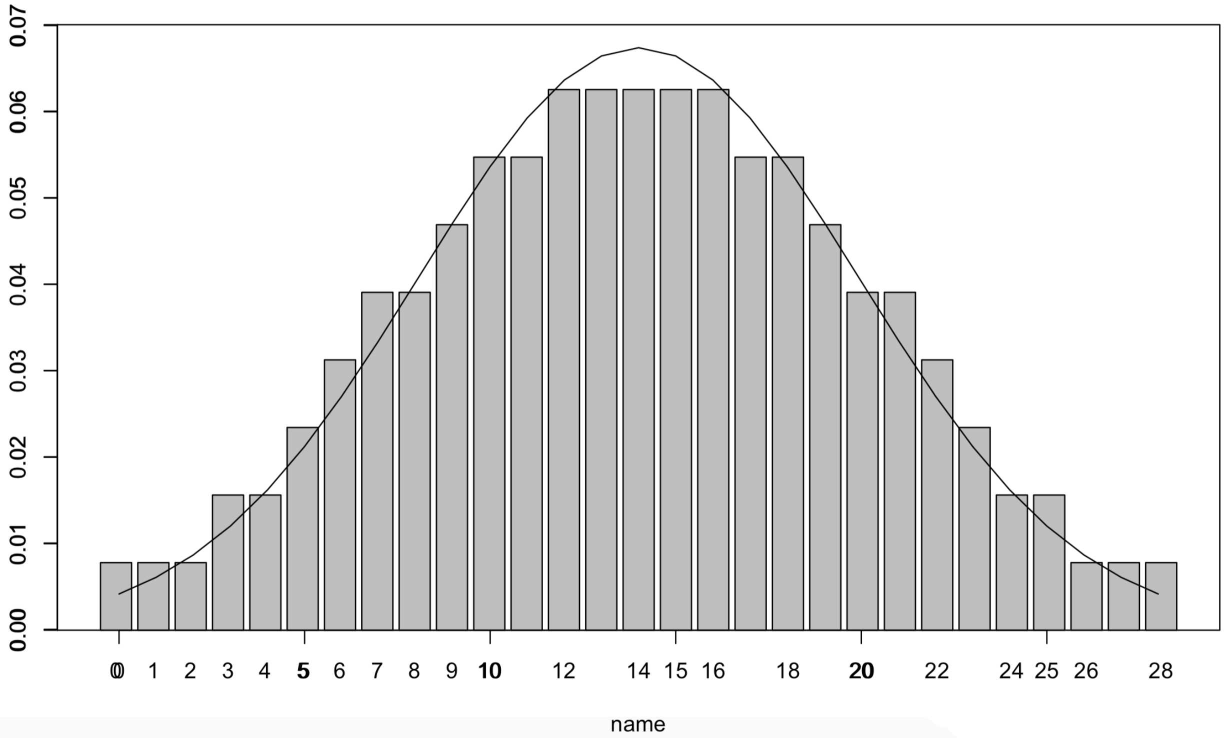
$\operatorname{Var}(S_{+})=\sum_{i=1}^{n}\left(s_i^2/4\right)=\frac{1}{4}\times\frac{n(n+1)(2n+1)}{6}=\frac{n(n+1)(2n+1)}{24}$

3.3 In Comment 3 on Example 3.2 we suggested that for a variety of reasons one should be cautious about extending inferences about heartbeat rates for female students to the population at large. What might some of these reasons be?

1. There may be differences between sexes.
2. There may be differences between ages within and between sexes.
3. Lifestyles of students may be very different from those of say, manual workers or office workers.
4. heriditary or environmental factors that may be very different between a student population and some other populations.

3.4 Using the data in Table 3.1 for the distribution of the Wilcoxon S when $n=7$, construct a bar chart like that in Figure 3.1 showing the probability function for S. Discuss the similarity, or lack of similarity, to a normal distribution probability density function.

The distribution is symmetric and unimodal, but discontinuities are quite evident and the general appearance is that the distribution is less peaked than the normal and lacks taper in the tails.




```
prob=c(1,1,1,2,2,3,4,5,5,6,7,7,8,8,8,8,8,7,7,6,5,5,4,3,2,2,1,1,1)
name=0:28
f=1/sqrt(2*pi)/sqrt(7*8*15/24)*exp(-(name-14)^2/2/(7*8*15/24))
barplot(prob/128,names.arg = name,ylim=c(0,0.07))
par(new=TRUE)
plot(name,f,xlim = c(-0.5,28.5),ylim = c(0.0025,0.0675),type='l')
```