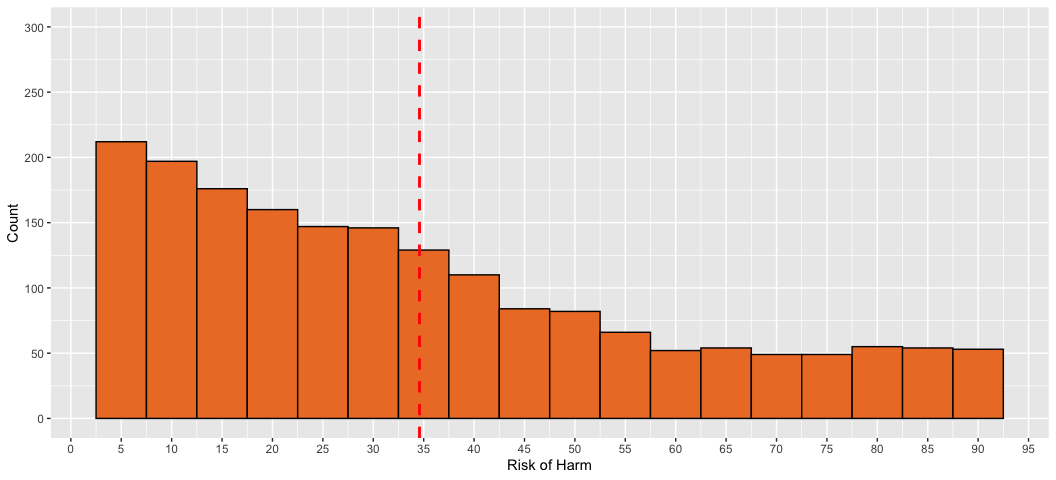
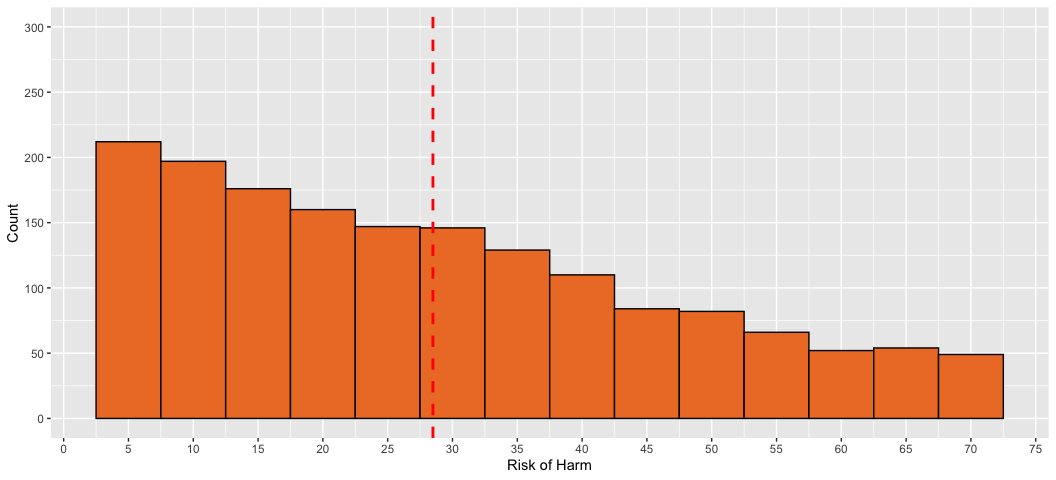
Here is the overall trial distribution for Yes responces for all subject as a fuction of risk and the same thing expect for a filter showing only yes responces for risk less than 70.

#Risk  
ggplot(data, aes(risk, fill = experimental\_treatment\_selected)) +  
 geom\_histogram(binwidth=5, colour="black", fill="#ee7c2e") +   
 geom\_vline(aes(xintercept=mean(risk, na.rm=T)),  
 color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Risk of Harm", breaks=seq(0, 100, 5))+  
 scale\_y\_continuous("Count", breaks=seq(0, 300, 50), limits=c(0, 300))



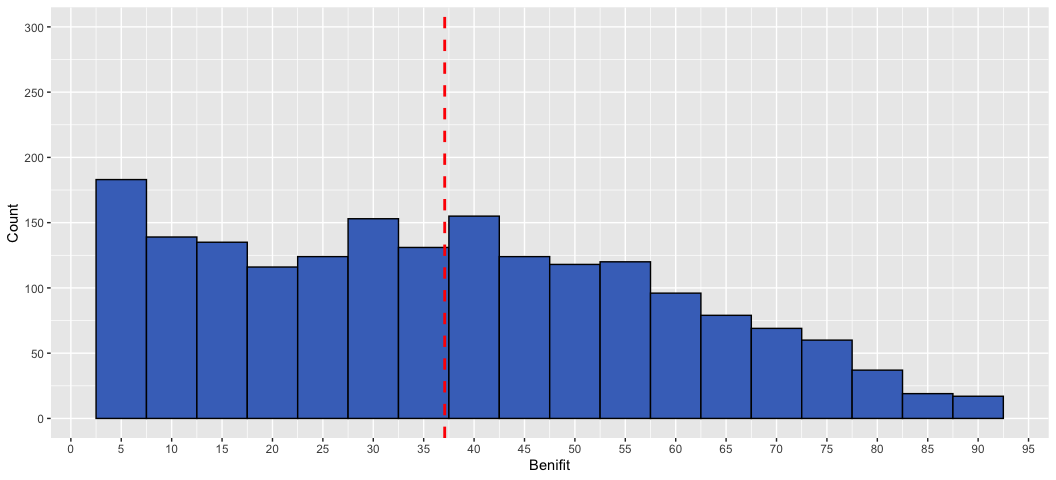
ggplot(data2, aes(risk, fill = experimental\_treatment\_selected)) +  
 geom\_histogram(binwidth=5, colour="black", fill="#ee7c2e") +   
 geom\_vline(aes(xintercept=mean(risk, na.rm=T)),  
 color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Risk of Harm", breaks=seq(0, 100, 5))+  
 scale\_y\_continuous("Count", breaks=seq(0, 300, 50), limits=c(0, 300))



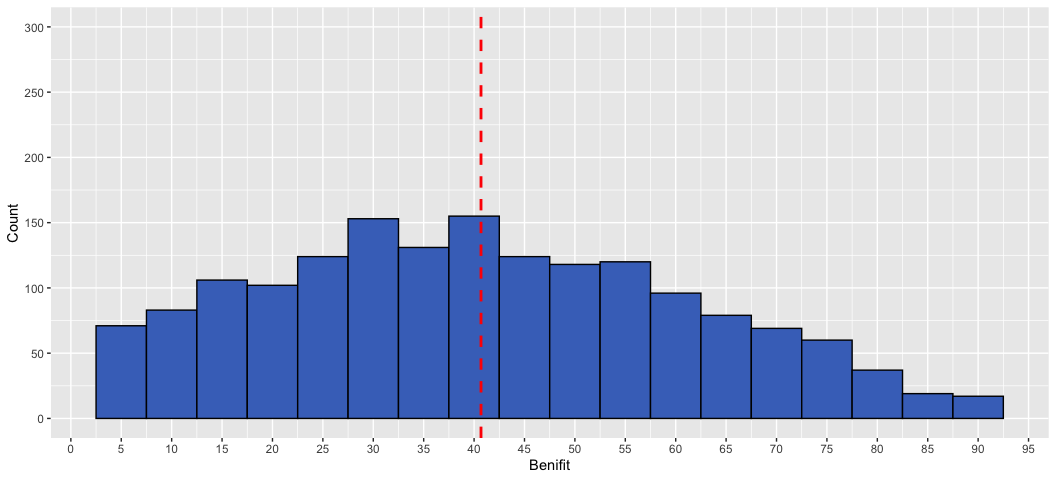
Of course we do not expect the shape to change here because we are only limiting the amount of observations we are seeing!

But here is the overall trial distribution for Yes responces for all subject as a fuction of gain and no effect with its contrasting filter of risk responces at 70.

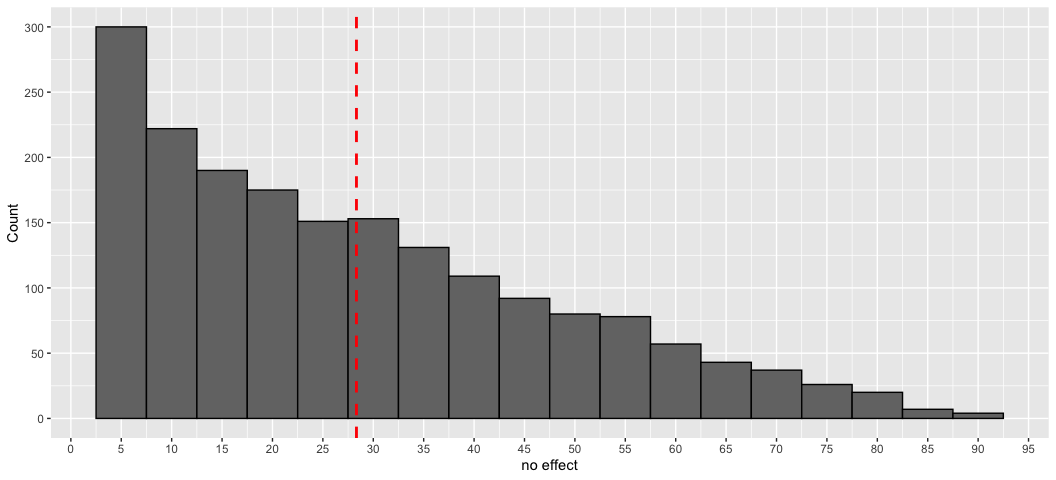
#Gain  
ggplot(data, aes(gain, fill = experimental\_treatment\_selected)) +  
 geom\_histogram(binwidth=5, colour="black", fill="#4673c3") +   
 geom\_vline(aes(xintercept=mean(gain, na.rm=T)),  
 color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Benifit", breaks=seq(0, 100, 5))+  
 scale\_y\_continuous("Count", breaks=seq(0, 300, 50), limits=c(0, 300))



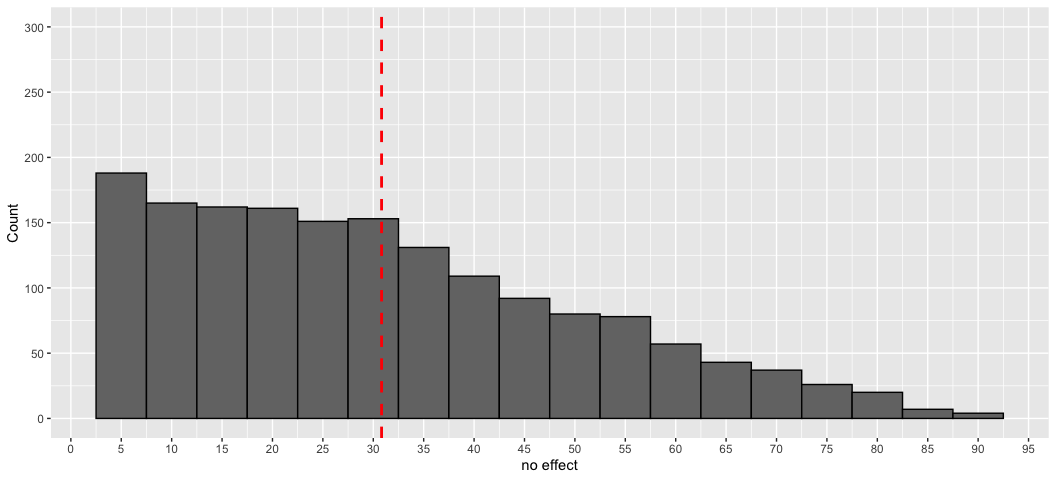
ggplot(data2, aes(gain, fill = experimental\_treatment\_selected)) +  
 geom\_histogram(binwidth=5, colour="black", fill="#4673c3") +   
 geom\_vline(aes(xintercept=mean(gain, na.rm=T)),  
 color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Benifit", breaks=seq(0, 100, 5))+  
 scale\_y\_continuous("Count", breaks=seq(0, 300, 50), limits=c(0, 300))



#No Effect  
ggplot(data, aes(no\_effect, fill = experimental\_treatment\_selected)) +   
 geom\_histogram(binwidth=5, colour="black", fill="#747474") +   
 geom\_vline(aes(xintercept=mean(no\_effect, na.rm=T)),  
 color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("no effect", breaks=seq(0, 100, 5))+  
 scale\_y\_continuous("Count", breaks=seq(0, 300, 50), limits=c(0, 300))



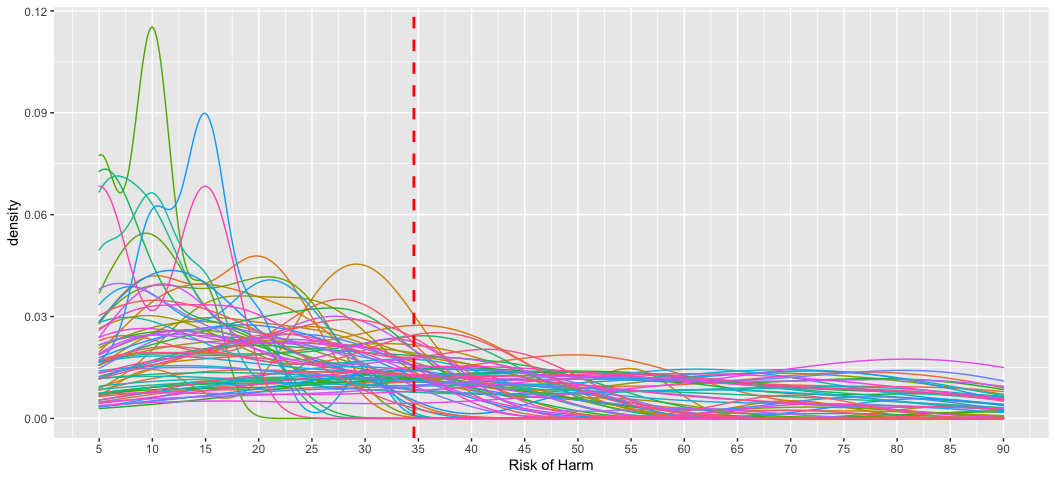
ggplot(data2, aes(no\_effect, fill = experimental\_treatment\_selected)) +   
 geom\_histogram(binwidth=5, colour="black", fill="#747474") +   
 geom\_vline(aes(xintercept=mean(no\_effect, na.rm=T)),  
 color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("no effect", breaks=seq(0, 100, 5))+  
 scale\_y\_continuous("Count", breaks=seq(0, 300, 50), limits=c(0, 300))

 We can see a significant change in our gain. This is because if we sample less risk, we also are sampling less trials with low amounts of gain and no effect. For example, if I never have a trial with 90 risk, I also shrink the amount of trials with 5 gain.

So over subjects, do we see a significant change in the distributions from this filtering?

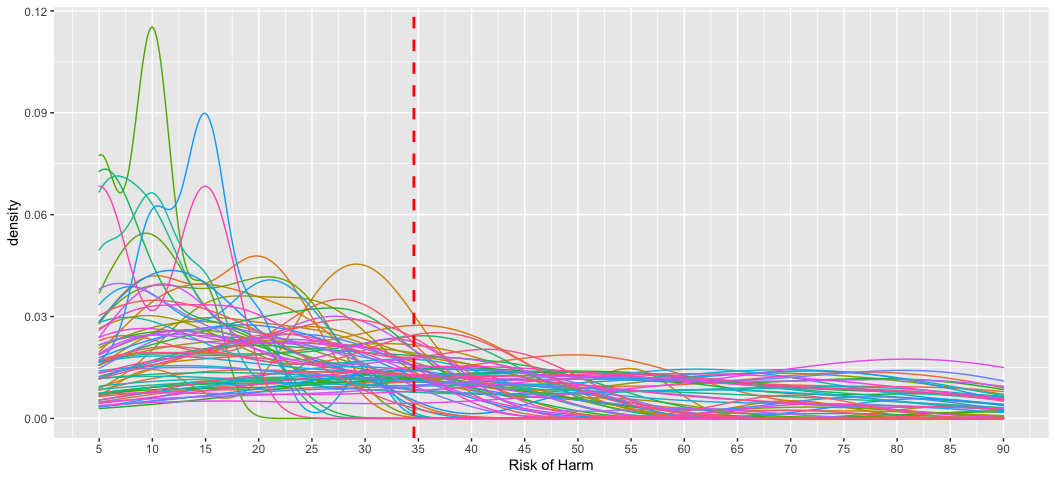
##Risk  
#Before filter  
ggplot(data, aes(x=risk, color=user\_id)) + geom\_density()+  
 geom\_vline(aes(xintercept=mean(risk, na.rm=T)), color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Risk of Harm", breaks=seq(0, 100, 5))+  
 theme(legend.position="none")

## Warning: Groups with fewer than two data points have been dropped.



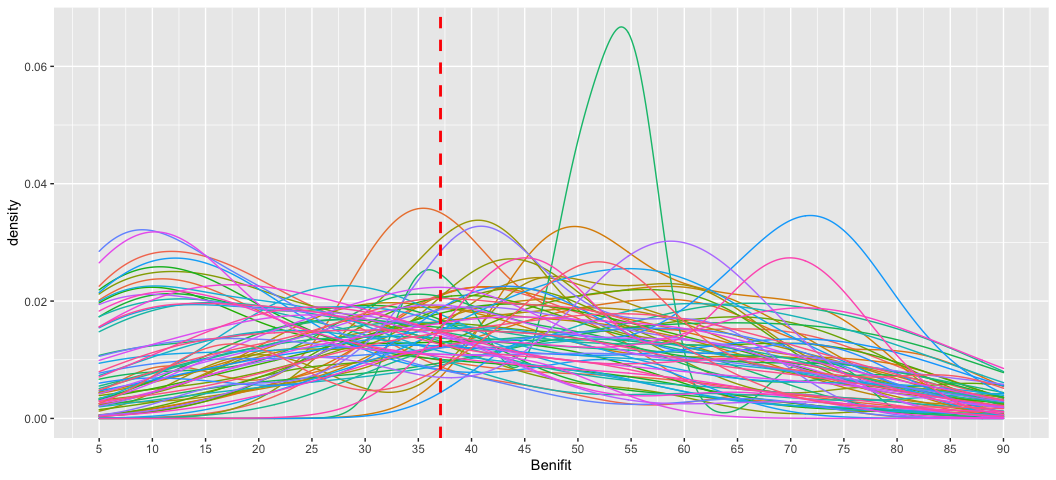
#After filter  
ggplot(data, aes(x=risk, color=user\_id)) + geom\_density()+  
 geom\_vline(aes(xintercept=mean(risk, na.rm=T)), color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Risk of Harm", breaks=seq(0, 100, 5))+  
 theme(legend.position="none")

## Warning: Groups with fewer than two data points have been dropped.



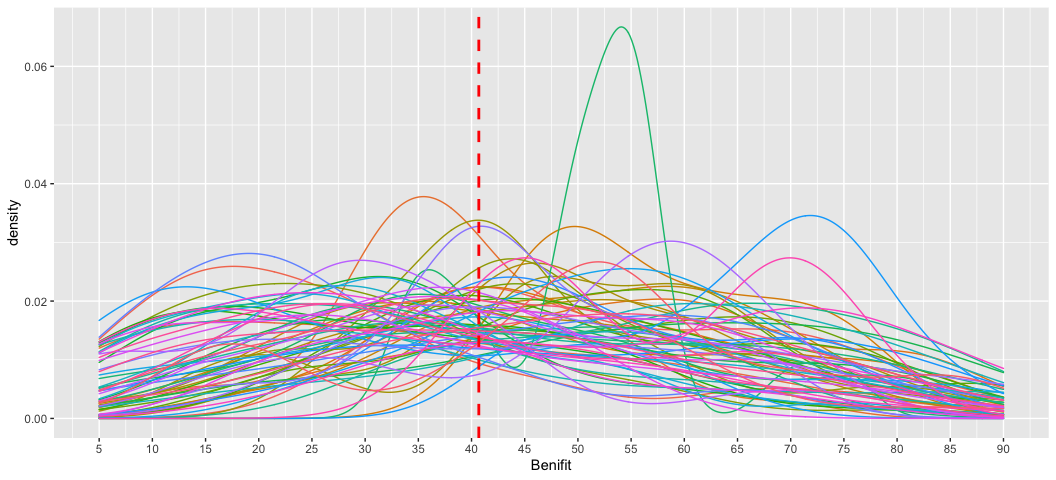
##Gain   
#Before filter  
ggplot(data, aes(x=gain, color=user\_id)) + geom\_density()+  
 geom\_vline(aes(xintercept=mean(gain, na.rm=T)), color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Benifit", breaks=seq(0, 100, 5))+  
 theme(legend.position="none")

## Warning: Groups with fewer than two data points have been dropped.



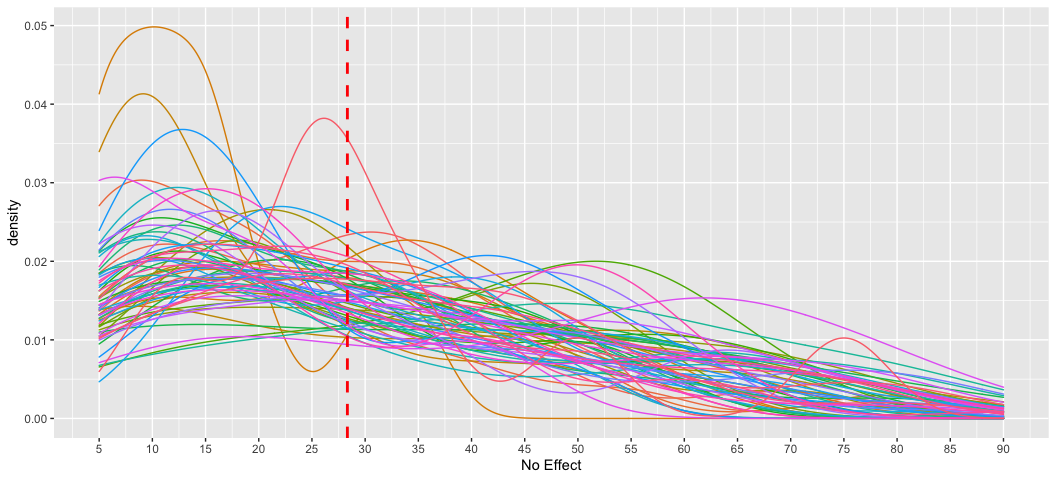
#After filter  
ggplot(data2, aes(x=gain, color=user\_id)) + geom\_density()+  
 geom\_vline(aes(xintercept=mean(gain, na.rm=T)), color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("Benifit", breaks=seq(0, 100, 5))+  
 theme(legend.position="none")

## Warning: Groups with fewer than two data points have been dropped.



#No effect  
#Before filter  
ggplot(data, aes(x=no\_effect, color=user\_id)) + geom\_density()+  
 geom\_vline(aes(xintercept=mean(no\_effect, na.rm=T)), color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("No Effect", breaks=seq(0, 100, 5))+  
 theme(legend.position="none")

## Warning: Groups with fewer than two data points have been dropped.



#After filter  
ggplot(data2, aes(x=no\_effect, color=user\_id)) + geom\_density()+  
 geom\_vline(aes(xintercept=mean(no\_effect, na.rm=T)), color="red", linetype="dashed", size=1)+  
 scale\_x\_continuous("No Effect", breaks=seq(0, 100, 5))+  
 theme(legend.position="none")

## Warning: Groups with fewer than two data points have been dropped.

