1. *Find the Marginal Distribution p(x)*

Lab Assignment 3 Dai Li

— Normal distribution $x \sim N(\mu, \sigma^2)$

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$$

Gamma distribution $x \sim Gamma(\alpha, \beta)$

$$f(x|\alpha,\beta) = \beta^{\alpha} \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x \geq 0, \alpha, \beta > 0$$

— Now, $x|\tau^2 \sim N(0, \frac{1}{\tau^2})$, $\tau^2 \sim Gamma(\frac{\nu}{2}, \frac{\nu}{2})$

So, $$\pi(\tau^2) = \left(\frac{\nu}{2}\right)^{\frac{\nu}{2}} \frac{1}{\Gamma(\frac{\nu}{2})} (\tau^2)^{(\frac{\nu}{2}-1)} e^{-\frac{\nu}{2}\tau^2}$$

and, $$f(x|0,\tau^2) = \frac{1}{\sqrt{2\pi(1/\tau^2)}} e^{-\frac{1}{2(1/\tau^2)}(x-0)^2}$$

$$= \frac{\sqrt{\tau^2}}{\sqrt{2\pi}} e^{(-\frac{\tau^2}{2}x^2)}$$

~~The marginal distribution of x, p(x)~~

$\cancel{p(x) = \int f(x|\tau^2)\, \pi(\tau^2)\, d\tau^2}$

— Consequently, The marginal distribution of x, p(x)

$\cancel{f(x, \tau^2|\nu)}$

can be derived from

$$p(x) = \int p(x|\tau^2)\, \pi(\tau^2)\, d\tau^2$$

$$= \int \frac{\sqrt{\tau^2}}{\sqrt{2\pi}} e^{(-\frac{\tau^2}{2}x^2)} \cdot \left(\frac{\nu}{2}\right)^{\frac{\nu}{2}} \frac{1}{\Gamma(\frac{\nu}{2})} (\tau^2)^{(\frac{\nu}{2}-1)} e^{-\frac{\nu}{2}\tau^2} d\tau^2$$

$$= \frac{(\nu/2)^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})\sqrt{2\pi}} \int_0^{\infty} (\tau^2)^{(\frac{\nu}{2}-\frac{1}{2})} \cdot e^{-\tau^2(\frac{1}{2}x^2 + \frac{1}{2}\nu)} d\tau^2$$

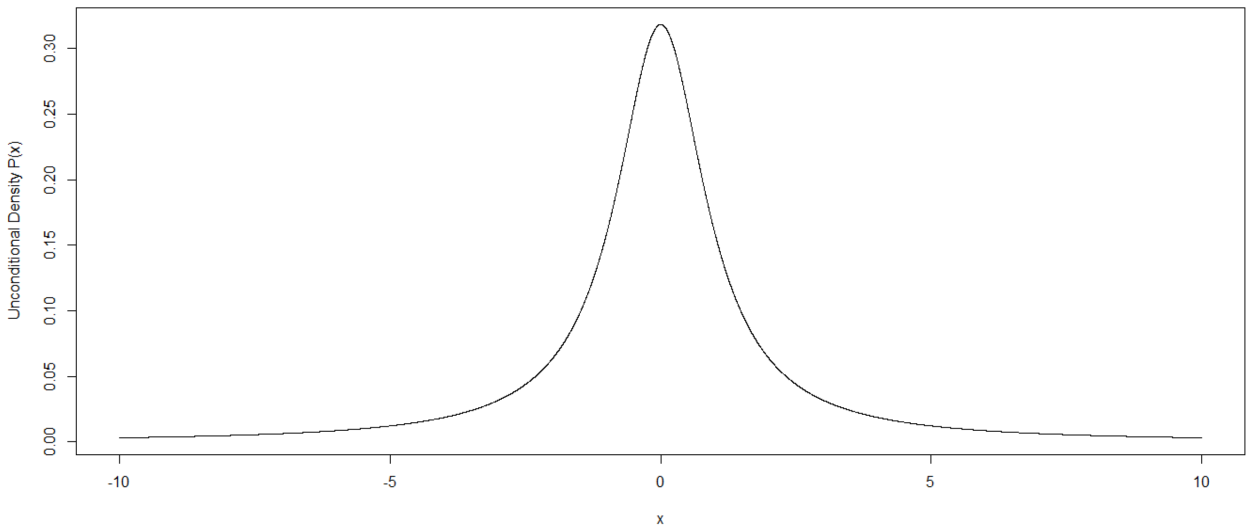
$$= \frac{(\nu/2)^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})\sqrt{2\pi}} \cdot \frac{\Gamma(\frac{\nu+1}{2})}{(\frac{1}{2}(x^2+\nu))^{\frac{\nu+1}{2}}}$$

$$= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$, which is

The pdf of students' t distribution.

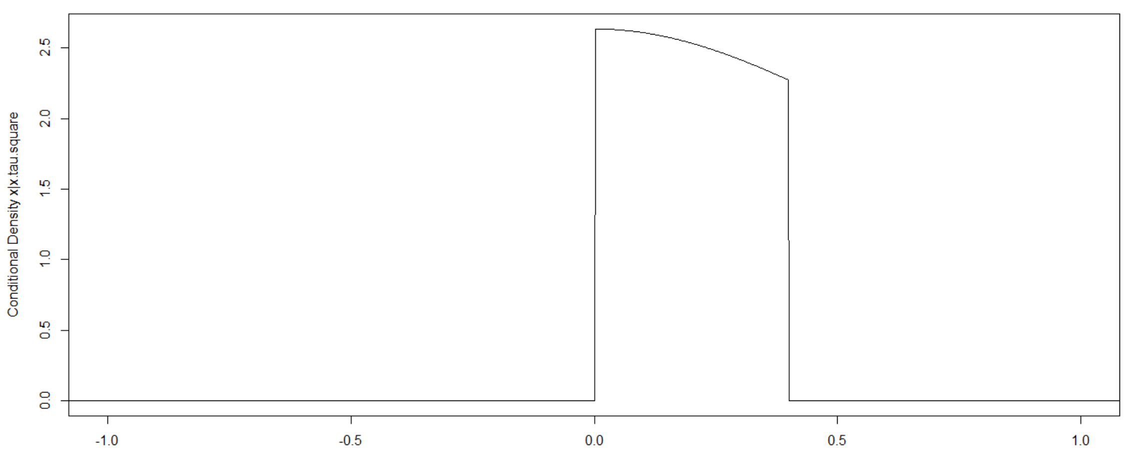
2. *Let nu=1. Get a sample of 10,000 from marginal distr. of x by drawing 10,000 tau^2's and then 10,000 x's given the tau^2's. Plot sample (either histogram or density is fine). Give two names for the actual marginal distribution p(x) when nu=1. Also, compute 2.5% & 97.5% percentile points of the distribution using the random samples and compare them to the theoretical values.*

The actual marginal distribution p(x) when nu=1 is plotted below:

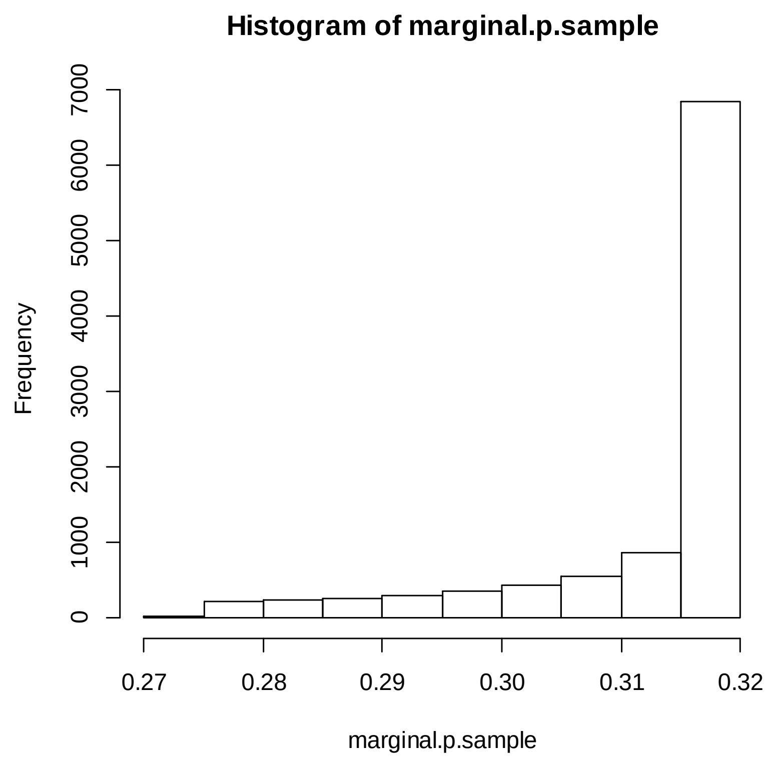


Which is a **<u>Students' T Distribution with parameter Nu=1</u>**. As, Nu=1, it is also a Cauchy Distribution with location parameter=0 and a scale parameter.

While using the 10000 sample points, we can draw out the theoretical conditional density of p(x|tau.square), which is plotted below:



Which is a Student's T distribution truncated to [min(x.tau.square.sample), max(x.tau.square.sample)], whose p(x)sample value histogram is also plotted below:



The 2.5% & 97.5% percentile points of theoretical p(x) are ***(-12.7062, 12.7062).*** While, the 2.5% & 97.5% percentile points of sampled p(x) are ***(0.009560445, 0.3877215).***

3*. Use Kolmogorov-Smirnov test (ks.test in R) to test whether your observed distribution is equal to a t(df=1). Report p-value. What is the conclusion of the test?*

Under null-hypothesis, p-value < 2.2e-16,

Using the sample points, D= 0.5853.

Conclusion of the test: R eject the hypothesis-***the observed distribution is not equal to a t(df=1).***

4*. Does the Central Limit Theorem hold for the mean of a sample from p(x) when nu=1? What about nu=2? nu=3? Why or why not? A quick explanation will do; an involved proof is NOT required.*

Given nu=1,2,or 3, there are finite value of X, so that there are finite value of mean E(x) and variance Var(x), and thus, the Central Limit Theorem hold for the sampled p(x).

Code:

```
theta=seq(0,100,length=10001);nu=1
tau.square <- dgamma(theta,(nu/2),(nu/2))
x=seq(-10,10,length=10001)
x.tau.square <- dnorm(0,1/tau.square)
marginal.p.unconditional <- dt(x,nu)
marginal.p.conditional <- dt(x,nu)*as.numeric(x > min(x.tau.square) & x <
max(x.tau.square))/(pt(max(x.tau.square),nu)-pt(min(x.tau.square),nu))


#draw 10000 samples
theta.sample <- runif(10000,0,1)
tau.square.sample <- dgamma(theta.sample,(nu/2),(nu/2))
x.tau.square.sample <- dnorm(0,1/tau.square.sample)
marginal.p.sample <- dt(x.tau.square.sample,nu)


#Plotting
par(mfrow = c(2,2))
plot(x,marginal.p.unconditional,xlim=c(-10,10),type="l",xlab="x",ylab="Unconditional Density P(x)")
plot(x,marginal.p.conditional,xlim=c(-1.0,1),type="l",xlab="x",ylab="Conditional Density x|x.tau.square")
hist(marginal.p.unconditional,breaks=10,plot=TRUE)
hist(marginal.p.sample,breaks=10,plot=TRUE)


#Calculating Credible Interval
truncated.inverse.cdf <- function(x,x.tau.square,nu){
F.a=pt(min(x.tau.square),nu)
F.b=pt(max(x.tau.square),nu)
z=qt((F.b-F.a)*x+F.a,nu)
return(z)
}
lower.theoretical <- qt(0.025,nu)
upper.theoretical <- qt(0.975,nu)
lower.actual <- truncated.inverse.cdf(.025,x.tau.square.sample,nu)
upper.actual <- truncated.inverse.cdf(.975,x.tau.square.sample,nu)

ks.test(marginal.p.sample,"pt",1)
```