

03-Variable Aleatoria Continua Unidimensional

Probabilidad Aplicada 3-602

Variable Aleatoria Continua

Una variable aleatoria continua es un valor numérico obtenido al azar, el cual resulta de hacer un experimento y que puede tomar infinitos valores. Esto significa que, conocidos dos valores consecutivos de la variable, siempre es posible encontrar otro valor intermedio entre ellos.

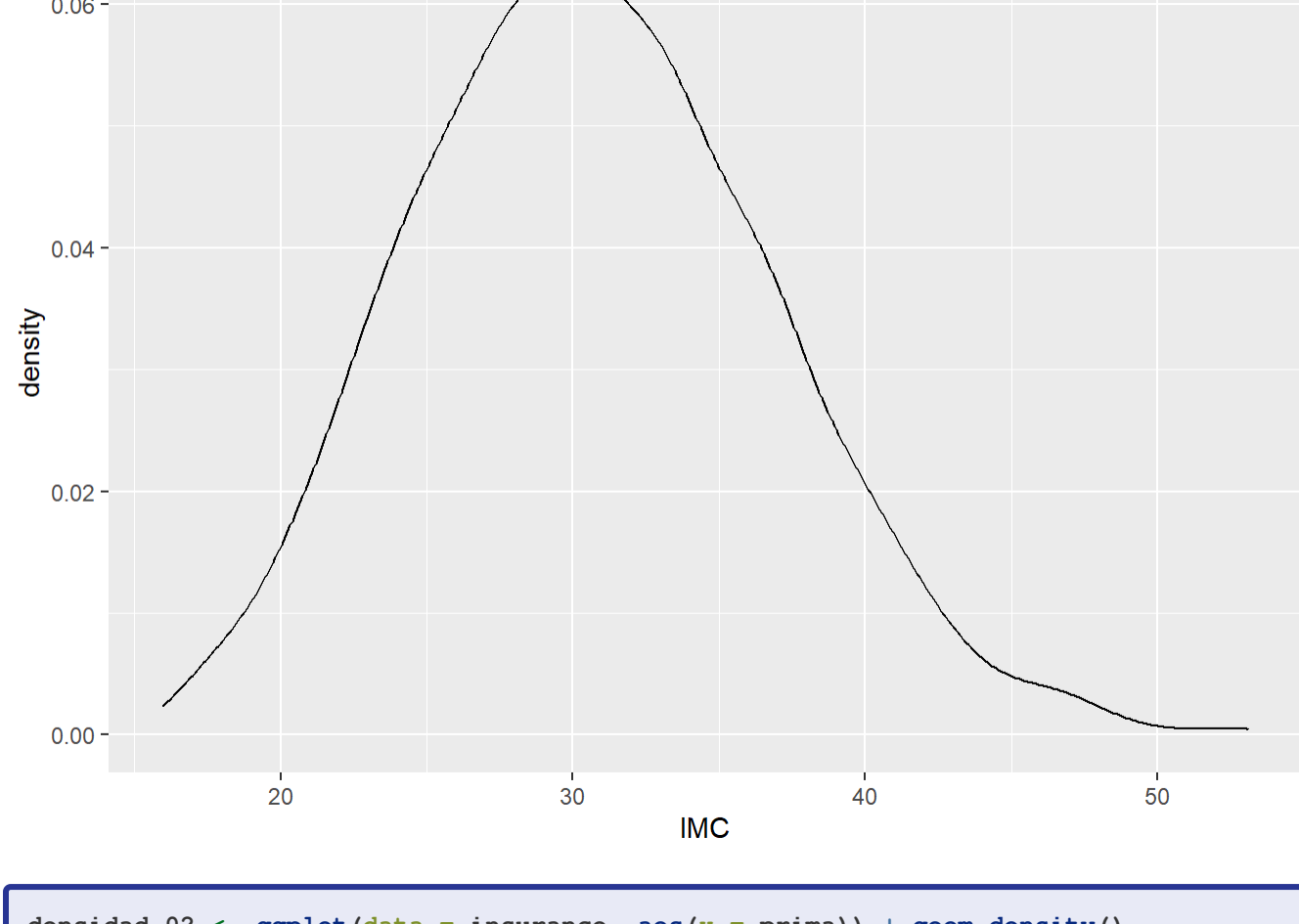
A modo de ejemplo, la estatura, el peso y la temperatura corporal de una persona son variables aleatorias continuas. La variable aleatoria continua no es contable mientras que la variable aleatoria discreta, si lo es. Por ejemplo el número de hijas en una familia, cuántos automóviles vende una agencia al cabo de un mes o cuántos .

Función de densidad

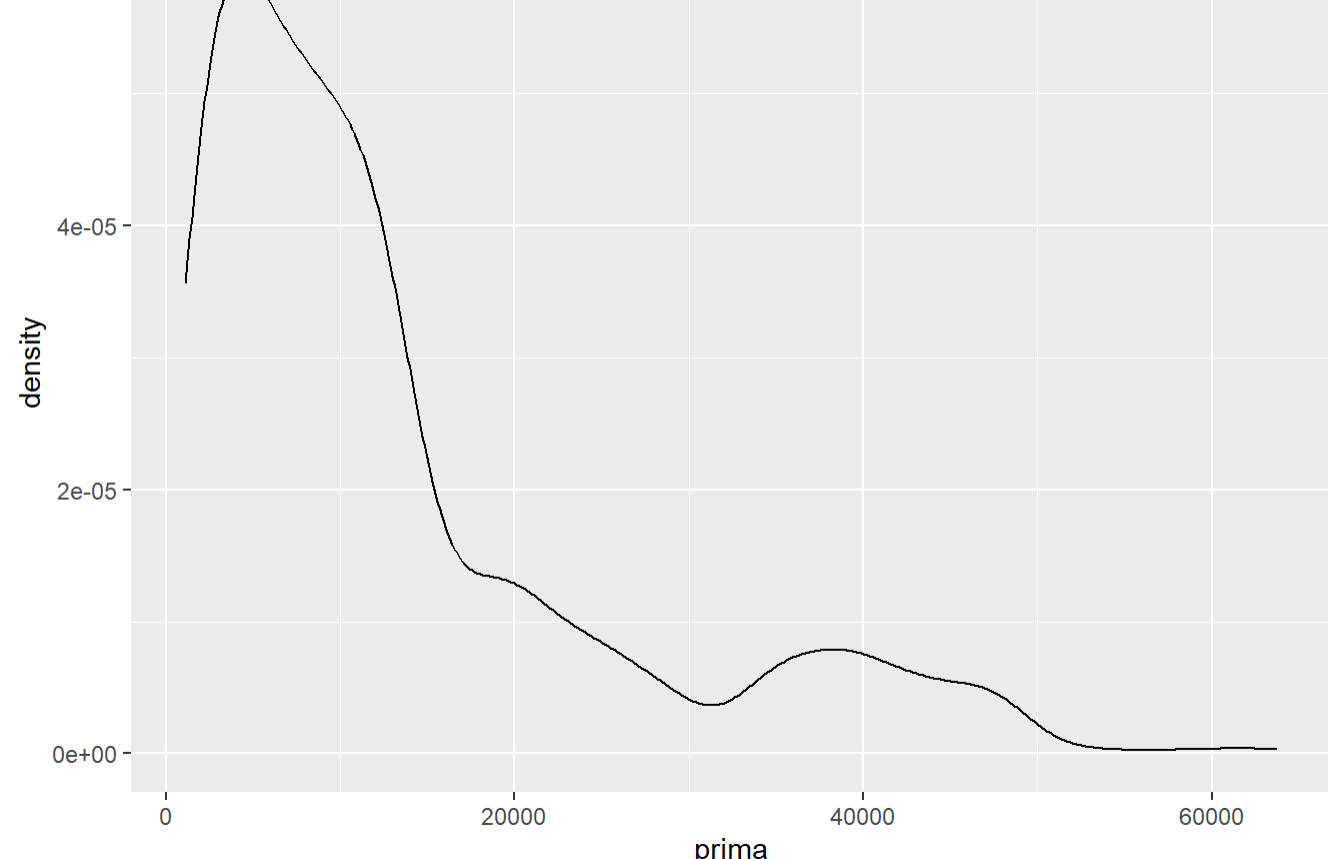
El histograma de frecuencias es muy útil para analizar la distribución de datos. Sin embargo depende del número de grupos o intervalos elegidos. Para resolver este problema, se utilizan las funciones de densidad, que es la curva resultante al considerar un número infinito de intervalos de longitud tendiente a cero que proporcionan infinitas barras (barras que tienden a líneas), cuyas alturas determinan la forma de la curva.

Gráficas con ggplot + geom_density()

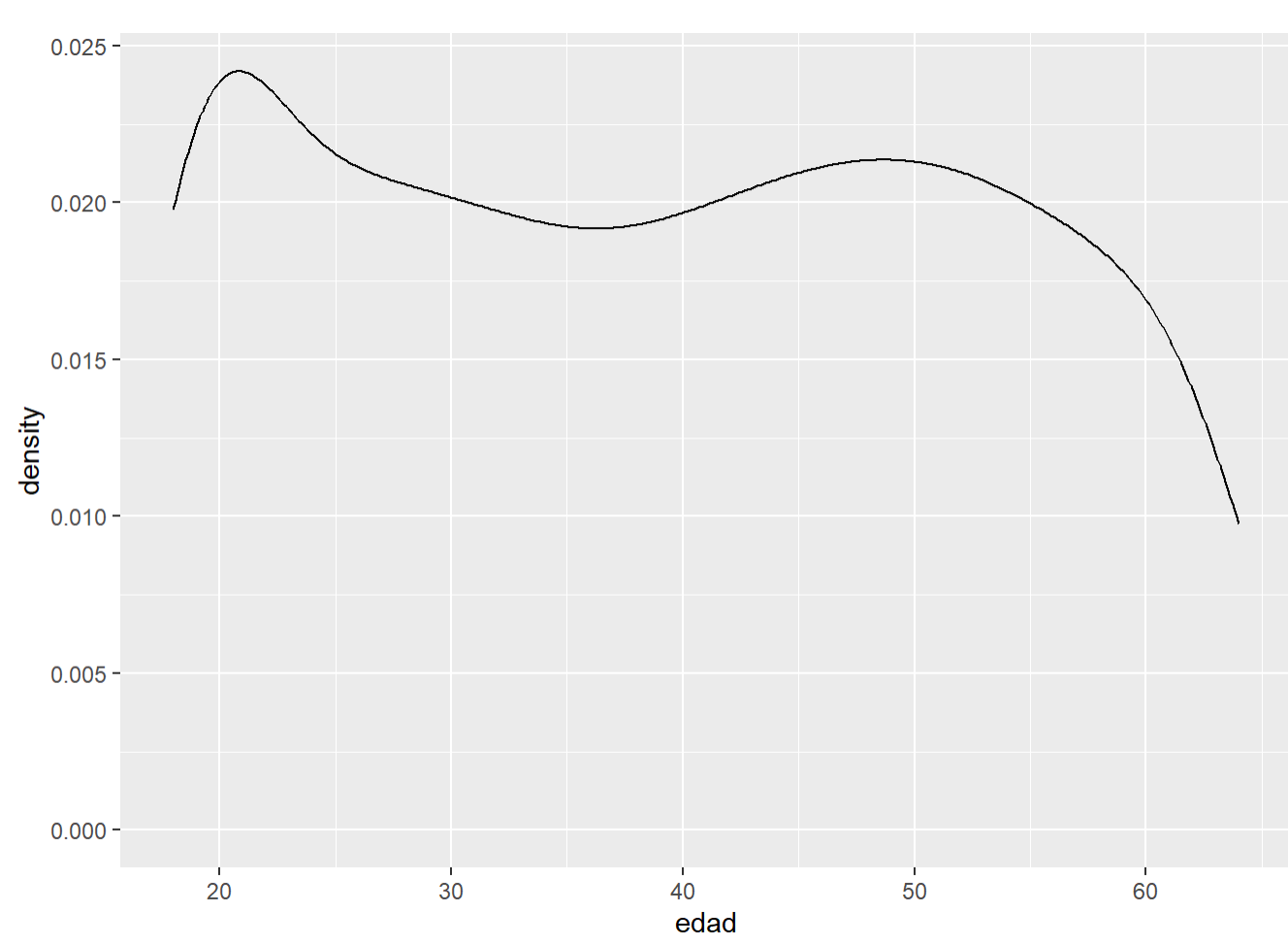
```
densidad_01 <- ggplot(data = insurance, aes(x = IMC)) + geom_density()
```



```
densidad_02 <- ggplot(data = insurance, aes(x = prima)) + geom_density()
```

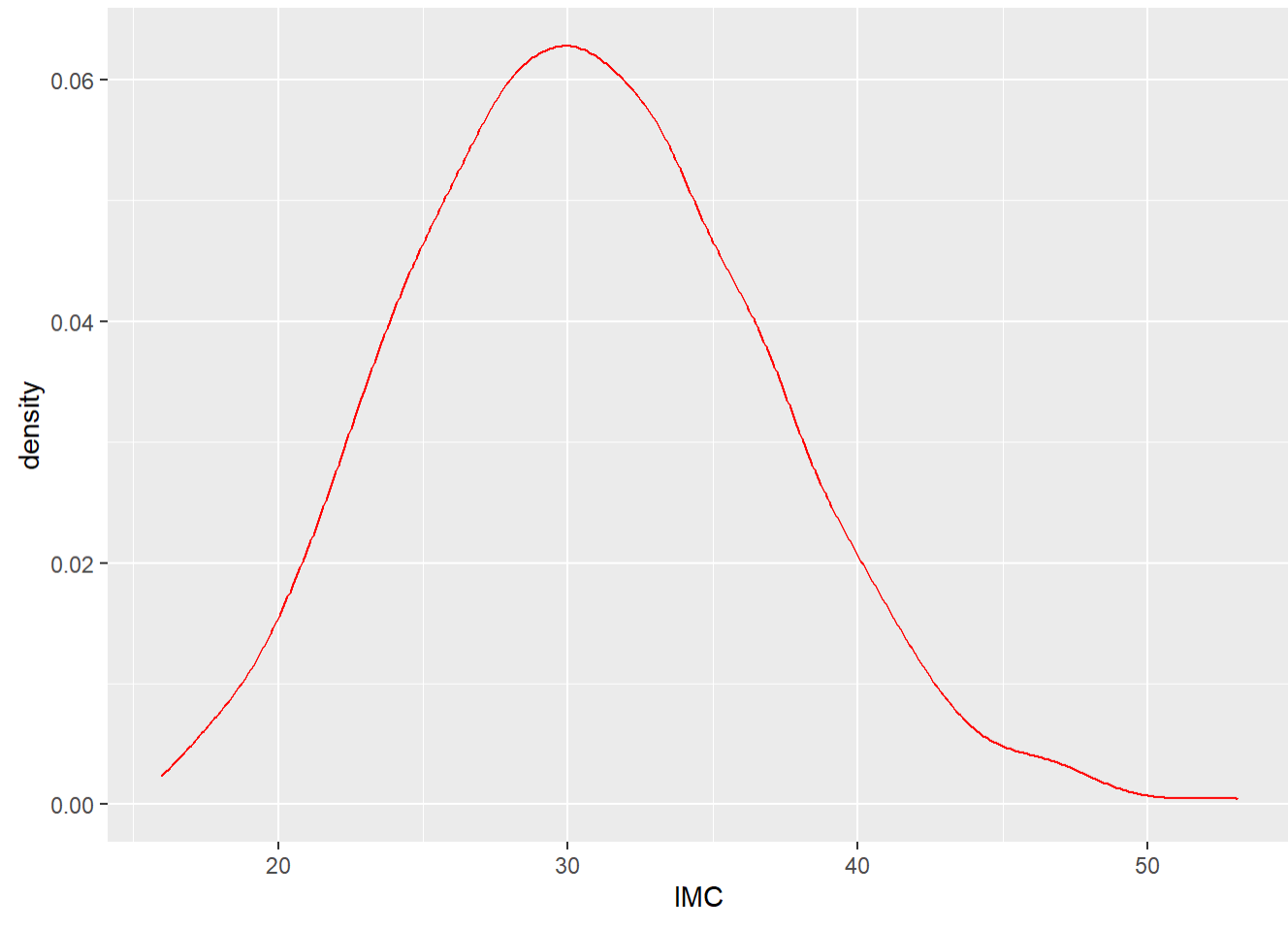


```
densidad_03 <- ggplot(data = insurance, aes(x = edad)) + geom_density()
```



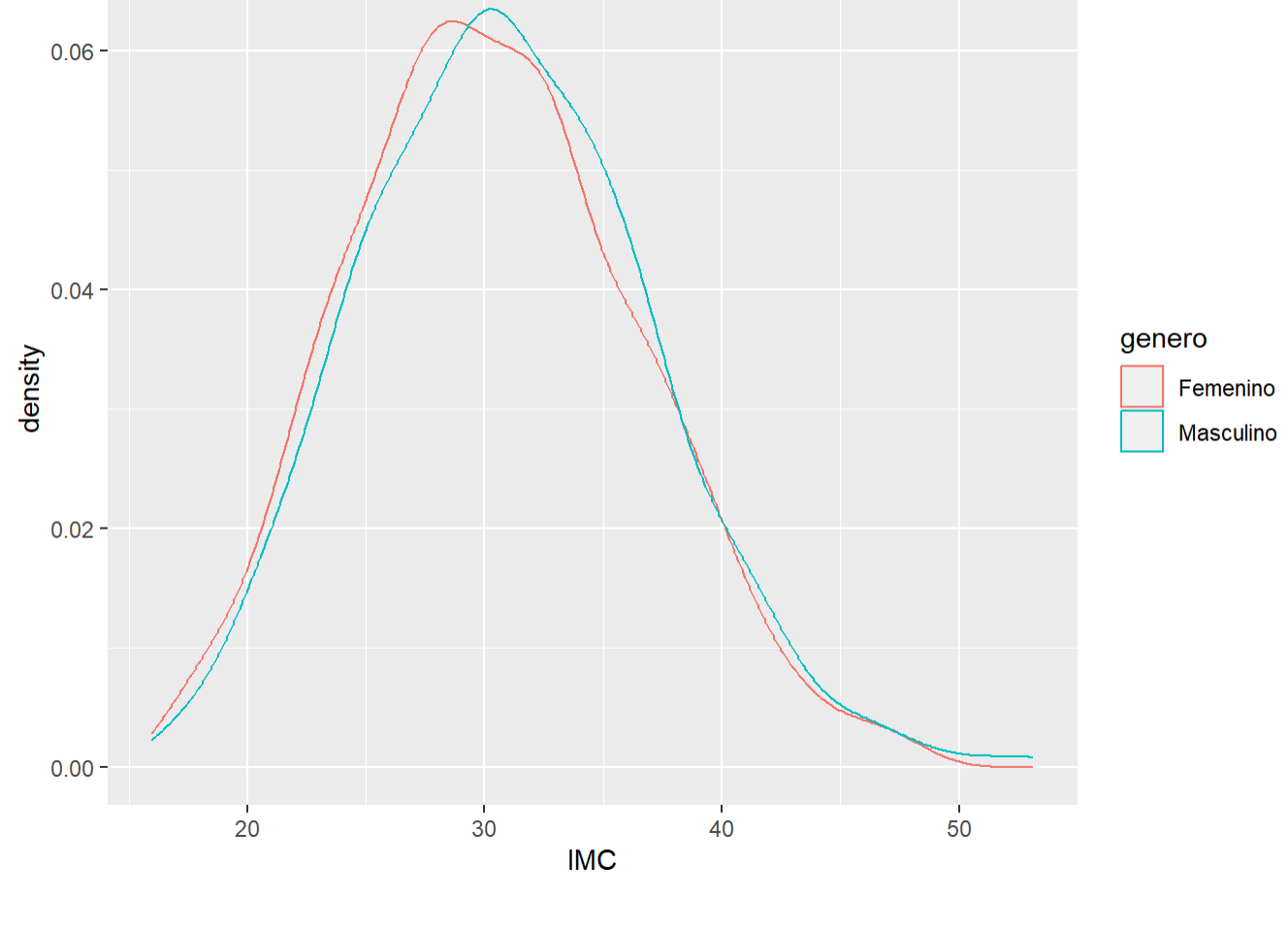
Mejorando las visualizaciones: Agregando color.

```
densidad_04 <- ggplot(data = insurance, aes(x = IMC)) + geom_density(color = "red")
```



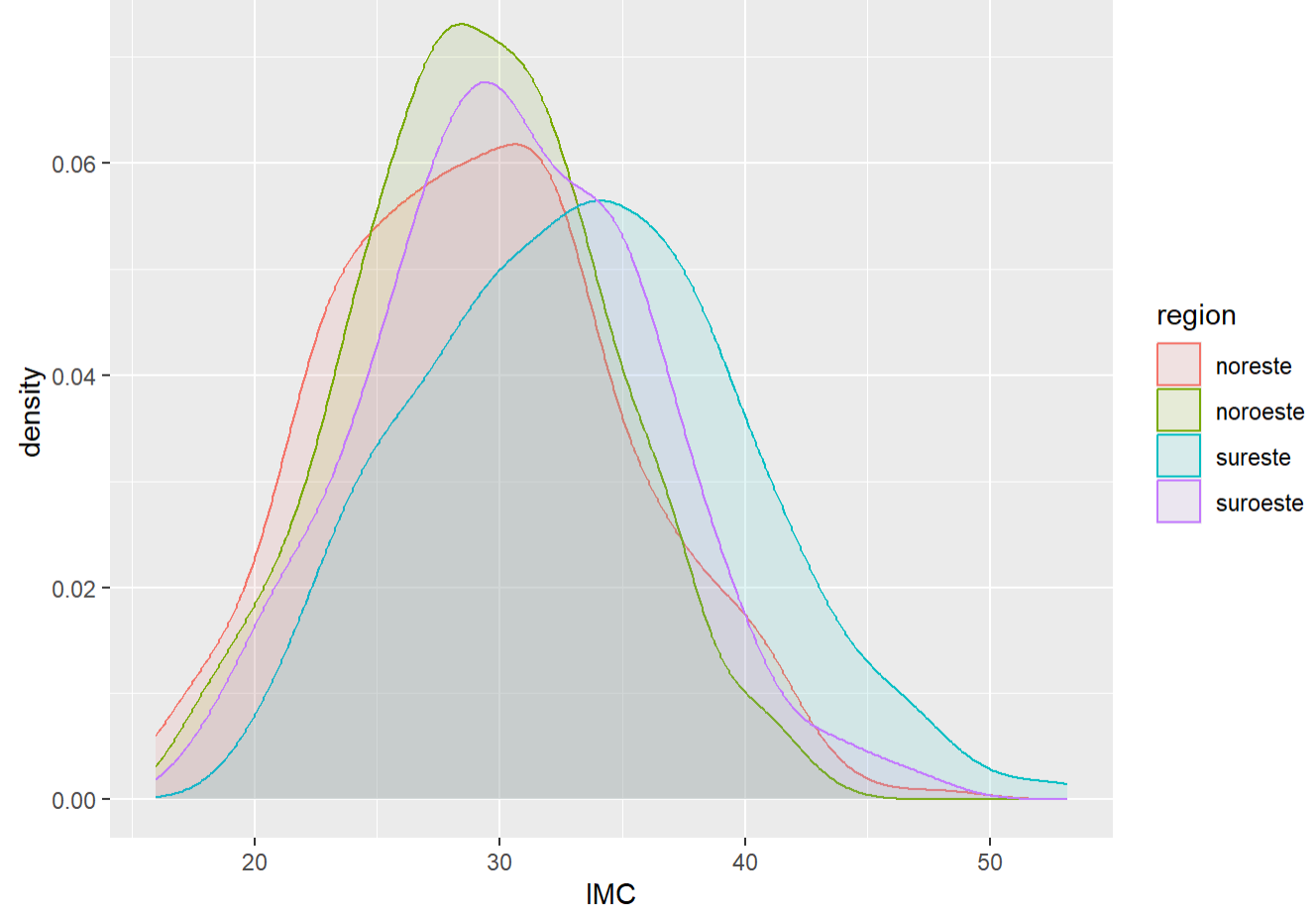
Mejorando las visualizaciones: Agregando filtrado por variable categórica.

```
densidad_05 <- ggplot(data = insurance, aes(x = IMC, colour = genero)) + geom_density()
```



Mejorando las visualizaciones: Agregando filtrado por variable categórica y rellenando.

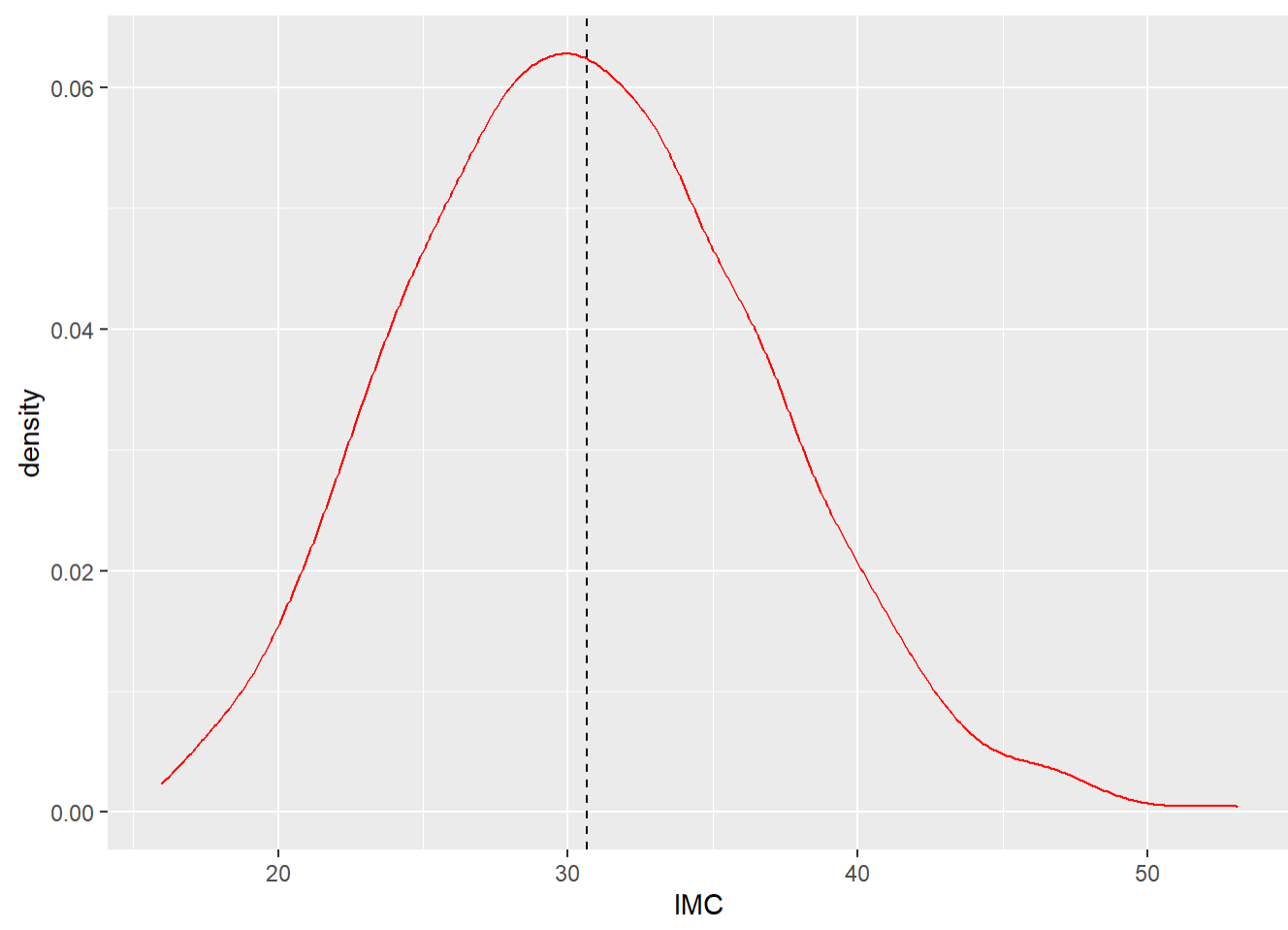
```
densidad_06 <- ggplot(data = insurance, aes(x = IMC, colour = region, fill = region)) +  
  geom_density(alpha = 0.1)
```



Mejorando las visualizaciones: Agregando linea de media.

```
densidad_07 <- ggplot(data = insurance, aes(x = IMC)) + geom_density(color = "red") +  
  geom_vline(aes(xintercept = mean(IMC)), linetype = "dashed")
```

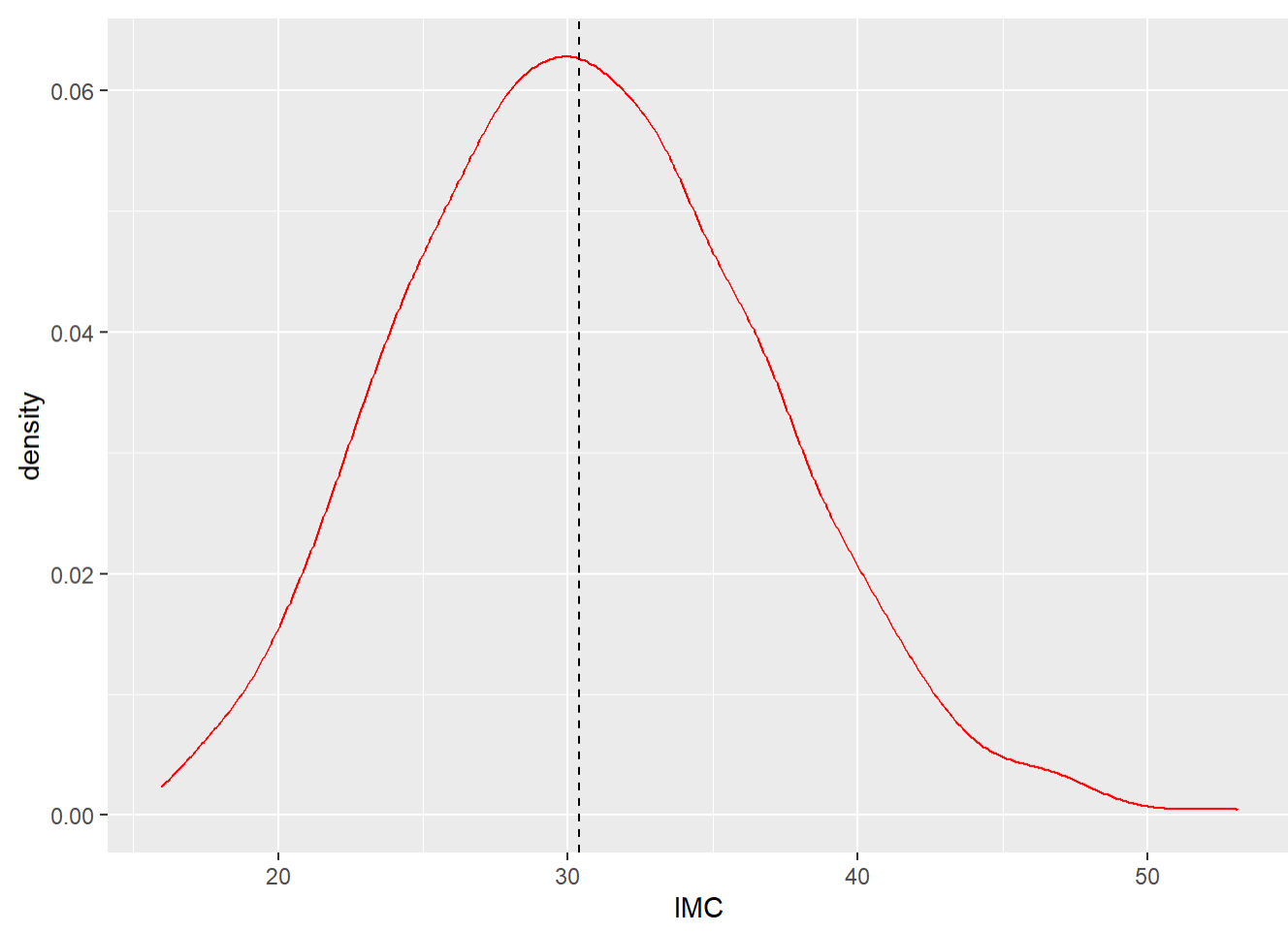
En este código, la función `mean(IMC)` proporciona la **media** de la variable IMC.



Mejorando las visualizaciones: Agregando linea de mediana.

```
densidad_08 <- ggplot(data = insurance, aes(x = IMC)) + geom_density(color = "red") +  
  geom_vline(aes(xintercept = median(IMC)), linetype = "dashed")
```

En este código, la función `median(IMC)` proporciona la **mediana** de la variable IMC.

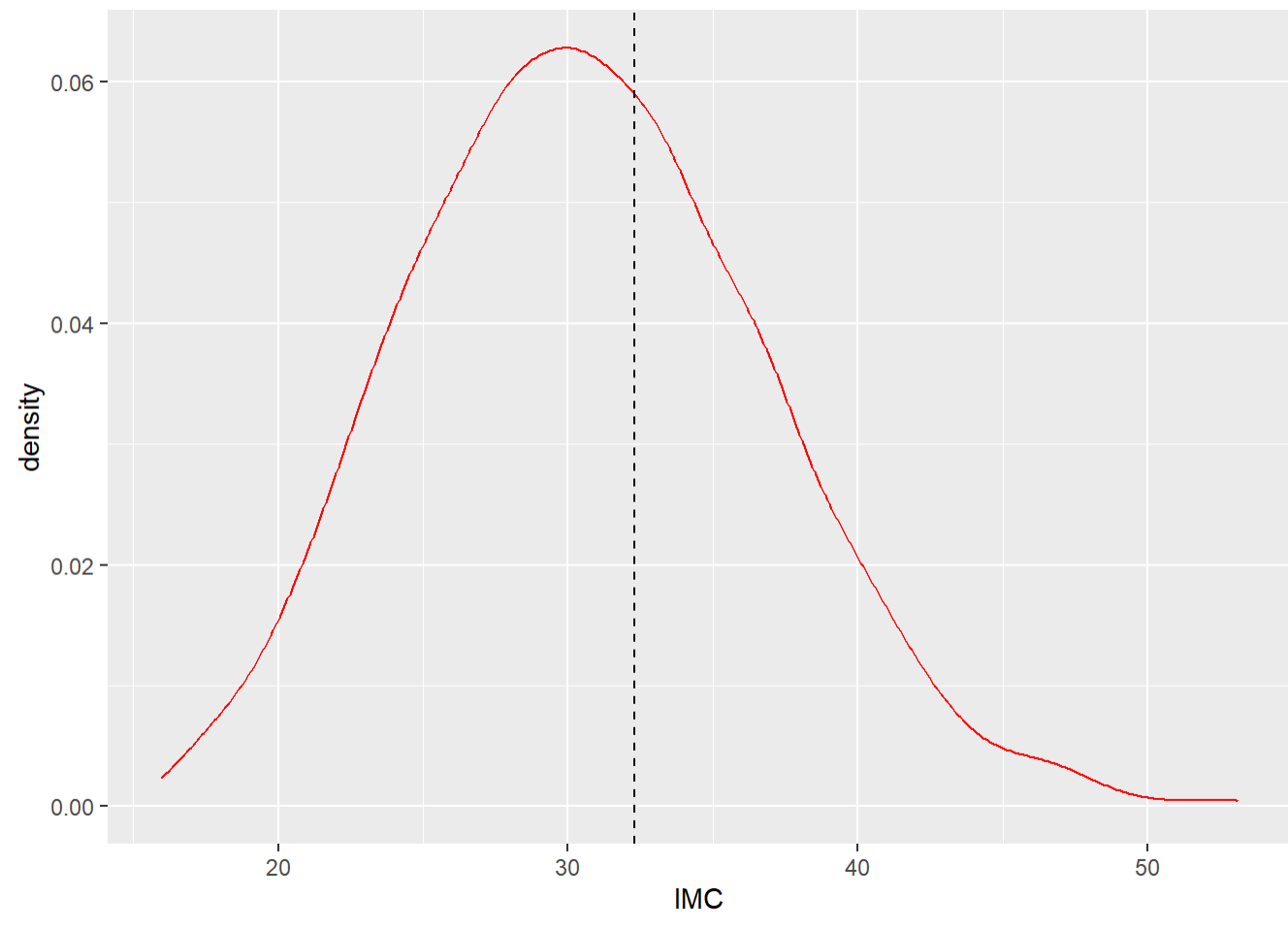


Mejorando las visualizaciones: Agregando linea de moda.

Para calcular la moda, es necesario cargar el paquete `modeest`.

```
densidad_09 <- ggplot(data = insurance, aes(x = IMC)) + geom_density(color = "red") +  
  geom_vline(aes(xintercept = mfv(IMC)), linetype = "dashed")
```

En este código, la función `mfv(IMC)` (mfv() de modeest) proporciona la **moda** de la variable IMC.



Las tres mediadas de tendencia central en un gráfico de densidad.

```
titulo <- "Gráfico de densidad de la variable IMC" #Definimos un título  
subtitulo <- paste("Media=",round(mean(insurance$IMC),1),  
  "Mediana=",median(insurance$IMC),  
  "Moda=",mfv(insurance$IMC)) #definimos un subtitulo que lleve los valores de las medi  
  
densidad_10 <- ggplot(data = insurance, aes(x = IMC)) + #creamos ggplot  
  ggtitle(titulo,subtitle = subtitulo)+#creamos una capa para el título/subtítulo  
  geom_density(color="red") + #creamos la capa con el gráfico de densidad  
  geom_vline(aes(xintercept = mean(IMC),color="Media"), linetype = "dashed",linewidth=1)+  
  geom_vline(aes(xintercept = median(IMC),color="Mediana"), linetype = "dashed",linewidth=1)+  
  geom_vline(aes(xintercept = mfv(IMC),color="Moda"), linetype = "dashed",linewidth=1)+  
  #creamos las capas con las rectas verticales de la tendencia central  
  scale_color_manual(names= "Medidas de tendencia central",  
    breaks=c("Media", "Mediana", "Moda"),  
    values =c("Media"="green", "Mediana"="blue", "Moda"="purple"))  
  #creamos una escala de color manual para poder referenciar.
```

