117- Cuando se condiciona sobre los valores de las variables independientes en la muestra, es claro que la distribución muestral de los estimadores de MCO depende de la distribución subyacente de los errores. Para hacer manejables las distribuciones de muestreo de los , ahora se supondrá que en la población el error no observado está distribuido normalmente. Esto se conoce como supuesto de normalidad.

El supuesto de normalidad supone los otros 5 supuestos. Los 6 supuestos se conocen como “supuestos del modelo lineal clásico”, y marcan que los estimadores insesgados de varianza mínima.

118- Aunque el teorema del límite central (TLC) puede satisfacerse en tales casos, la aproximación normal puede ser mala dependiendo de cuántos factores aparezcan en u y de qué tan diferentes sean sus distribuciones. Otro problema serio con el argumento del TLC es que éste supone que todos los factores no observados afectan a y por separado en forma aditiva. Nada garantiza que esto sea así. Si u es una complicada función de los factores no observados, entonces el argumento del TLC no puede emplearse.

119- Ejemplo de salarios (salario mínimo, salario = 0). Evidencias empíricas anteriores indican que la normalidad no es un buen supuesto en el caso del salario. Con frecuencia, usando una transformación, en especial empleando el logaritmo, se obtiene una distribución que es cercana a la normal.

Diagrama

Descripción generada automáticamente

El que no haya normalidad en los errores no es un problema serio cuando los tamaños de muestra son grandes. Por ahora, sólo se hace el supuesto de normalidad.

Estos 6 supuestos permiten el test de hipótesis.



123- Pruebas contra alternativas de una cola. Nivel de significancia del 5%.

124- Para grados de libertad mayor a 120, pueden emplearse los valores críticos de la normal estándar.

Histograma

Descripción generada automáticamente

125- Ejemplo salarios – educ.

127- Uso de R2 para elegir la mejor forma funcional.

130-

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

133- ¿cuál es el menor nivel de significancia al que se habría rechazado la hipótesis nula? Este nivel se conoce como el valor-p. El valor-p es el nivel de significancia de la prueba cuando se usa el valor del estadístico de prueba.

134- Pero obtener el valor-p unilateral es sencillo: simplemente se divide el valor-p de dos colas entre 2.

135- significancia económica o significancia práctica de esa variable está relacionada con la magnitud (y signo) de estimado.

136- El ejemplo anterior evidencia que cuando se trabaja con muestras grandes, además de mirar al estadístico t, es muy importante interpretar la magnitud de los coeficientes. Cuando se tienen tamaños de muestra grandes, los parámetros pueden estimarse de manera muy precisa: los errores estándar suelen ser bastante pequeños en relación con los coeficientes estimados, lo que en general da como resultado significancia estadística.

140- Hay que recordar que un intervalo de confianza sólo puede ser tan bueno como los supues tos subyacentes empleados para construirlo. Si se han omitido factores importantes que estén correlacionados con las variables explicativas, las estimaciones de los coeficientes no son con fiables: MCO es sesgado. Si existe heterocedasticidad —como por ejemplo, si que la varianza de log(rd) depende de alguna de las variables explicativas— entonces el error estándar no es válido como estimación de() (como se analizó en la sección 3.4) y el intervalo de confianza calculado usando estos errores estándar no será en verdad un IC de 95%. Se ha usado también el supuesto de normalidad en los errores para obtener estos IC, pero, como se verá en el capítulo 5, esto no es tan importante en aplicaciones en las que se cuenta con cientos de observaciones.

140- Pruebas de hipótesis para los coeficientes de dos variables, o combinaciones lineales (educación carrera universitaria o terciaria). El mayor problema es la obtención del error estándar.

Imagen que contiene Icono

Descripción generada automáticamente

143- Pruebas para restricciones lineales múltiples: la prueba F (o prueba para las restricciones de exclusión). Prueba de hipótesis múltiple o conjunta.

145- Modelo restringido (sin las variables en cuestión) y no restringido (con las variables en cuestión).

Diagrama

Descripción generada automáticamente

Para emplear el estadístico F es necesario conocer su distribución de muestreo bajo la hipó tesis nula para de esta manera poder elegir los valores críticos y la regla de rechazo. Se puede demostrar que, bajo H0 (y suponiendo que se satisfacen los supuestos del MLC), F está distribuida como una variable aleatoria F con (q,n k 1) grados de libertad.

147- Estadísticamente significativas conjuntamente, o conjuntamente significativas. El estadístico F suele ser útil para probar la exclusión de un grupo de variables cuando esas variables están fuertemente correlacionadas.

149- En una prueba es posible que se agrupen variables no significativas con una variable significativa y se concluya que todo el conjunto de variables es conjuntamente no significativo. (Estos conflictos entre una prueba t y una prueba conjunta F proporcionan otro ejemplo de por qué no debe decirse que se “acepta” la hipótesis nula; sólo debe decirse que no se puede rechazar.) El estadístico F sirve para detectar si un conjunto de coeficientes es distinto de cero, pero nunca será la mejor prue ba para determinar si un solo coeficiente es distinto de cero. Para probar una sola hipótesis la prueba t es más adecuada.

Alguien que desee concluir que la raza no es un factor significativo podría reportar algo así como “Las variables raza y edad fueron incluidas en la ecuación, pero resultaron conjuntamente no significativas al nivel de 5%”. Tal vez una cuidadosa revisión evite esta clase de conclusiones equivocadas, pero el lector debe ser advertido de que tales resultados pueden encontrarse. Con frecuencia, cuando una variable estadísticamente muy significativa se prueba junto con otro grupo de variables, éste es conjuntamente significativo. En tales casos no hay ninguna inconsistencia lógica al rechazar ambas hipótesis nulas.