Probabilidad y Estadística para Bachillerato, CIMAT, ene-jun 2018
(9 de abril 2018)

Del libro de Wackerly, Mendenhall y Scheaffer::

1. Prob. 3.119.- Se reparten cartas al azar y sin reemplazo de una baraja de 52 cartas. ¿Cuál es la probabilidad de que el segundo rey se reparta en la quinta carta?

   En las primeras 4 sólo puede haber un rey y, en la quinta, debe aparecer el otro rey.

$$P(\text{2o. rey en 5a. extracción}) = P(\text{1 rey en 4 extracciones}) \times P(\text{rey}|\text{1 rey en 4 extracciones})$$
$$= \frac{\binom{4}{1}\binom{48}{3}}{\binom{52}{4}} \times \frac{3}{48} = 0.01597$$

```
(4*48*47*46*3)/(26*51*25*49*48)                         # 0.01597193
dhyper(x=1,m=4,n=52-4,k=4) * dhyper(x=1,m=3,n=48-3,k=1) # 0.01597193
```

2. Ejemplo 3.17.- Un producto industrial se envía en lotes de 20. Es costoso realizar pruebas para determinar si un artículo es defectuoso y, por tanto, el fabricante muestrea su producción en lugar de usar un plan de inspección al 100%. Un plan de muestreo, construido para minimizar el número de piezas defectuosas enviadas a los clientes, exige muestrear cinco artículos de cada lote y rechazar el lote si se observa más de una pieza defectuosa. (Si el lote es rechazado, cada artículo del mismo se prueba posteriormente.) Si un lote contiene cuatro piezas defectuosas, ¿cuál es la probabilidad de que sea rechazado? ¿Cuál es el número esperado de piezas defectuosas en la muestra de tamaño 5? ¿Cuál es la varianza del número de piezas defectuosas de la muestra de tamaño 5?

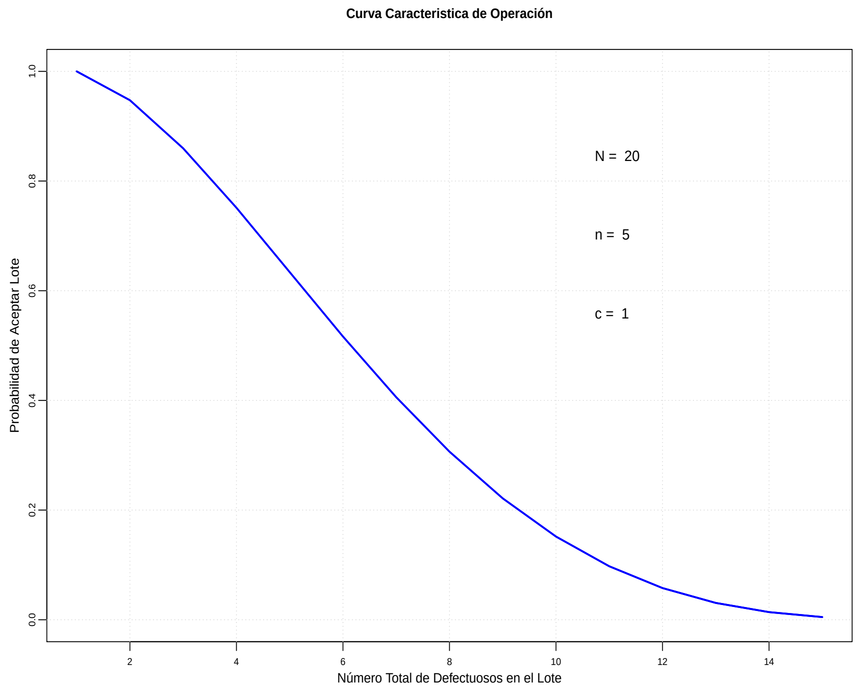
```
N = 20
n = 5
K = 4
1 - (choose(K,0)*choose(N-K,5)+ choose(K,1)*choose(N-K,4))/choose(N,n) #  0.24871
1 - phyper(1,K,N-K,n)   #  0.24871
```

   La media y varianza del número de defectuosos en la muestra de tamaño 5 son:

$$\mathsf{E}(X) = n\,\frac{K}{N} = 5\,\frac{4}{20} = 1 \qquad \mathsf{Var}(X) = n\,\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1} = 5\,\frac{4}{20}\left(1-\frac{4}{20}\right)\frac{20-5}{20-1} = 0.6316$$

   Podemos darnos una mejor idea de las propiedades de ese plan de aceptación, calculando la probabilidad de aceptar el lote como función del número real de defectuosos:

```
N   = 20
n   = 5
K   = 1:15  # Número real de defecuosos en el lote
c   = 1     # Aceptar si No. def. <= c
acp = phyper(c, K, N-K, n)
plot(K,acp,type="l",mgp=c(1.5,.5,0), xlab="Número Total de Defectuosos en el Lote",
     ylab="Probabilidad de Aceptar Lote",lwd=2,col="blue",
     cex.lab=.9, cex.axis=.7, ylim=c(0,1),
     main="Curva Caracteristica de Operación",cex.main=.9)
grid()
lines(K,acp,lwd=2,col="blue")
legend(10,.99,bty="n",legend=c(paste("N = ",N),paste("n = ",n),paste("c = ",c)))
```

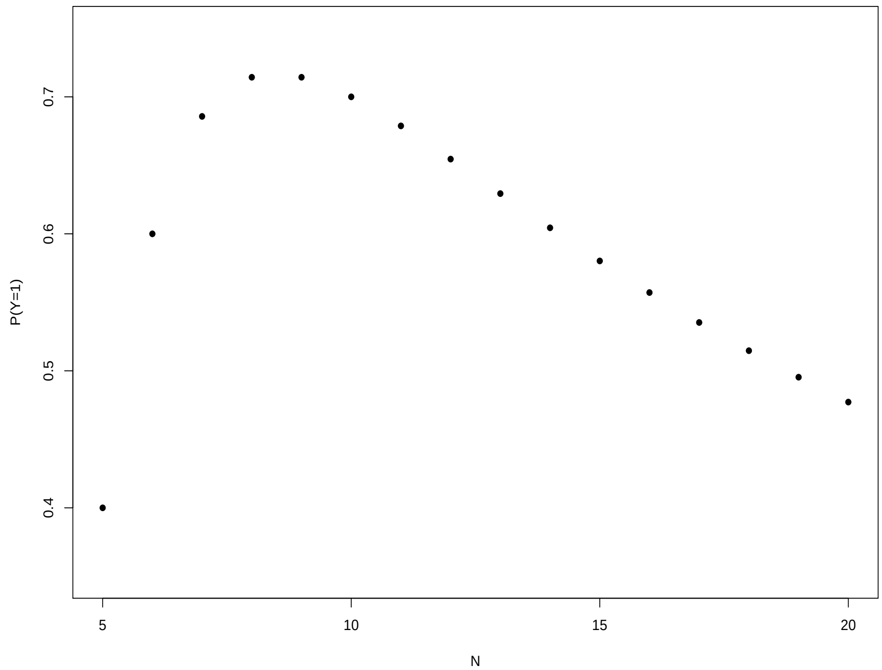


3. Prob. 3.120.- Los tamaños de poblaciones de animales se calculan en ocasiones con el método de capturar, marcar y recapturar. En este método se capturan k animales, se marcan y luego se sueltan en la población. Cierto tiempo después se capturan n animales y se observa Y, el número de animales marcados de entre los n. Las probabilidades asociadas con Y son una función de N, el número de animales de la población, de modo que el valor observado de Y contiene información sobre esta N desconocida. Suponga que $k = 4$ animales son marcados y luego soltados. Una muestra de $n = 3$ animales se selecciona entonces al azar de entre la misma población. Encuentre $P(Y = 1)$ como función de N. ¿Qué valor de N maximizará $P(Y = 1)$?

$$P(Y=1) = \frac{\binom{k}{1}\binom{N-k}{n-1}}{\binom{N}{n}} = 12 \times \frac{(N-4)(N-3)}{N(N-1)(N-2)}$$

La probabilidad de que $Y = 1$ es maximizada cuando $N = 8$ ó $N = 9$.

```
N   = 5:20
pr1 = 12*(N-4)*(N-3) / (N*(N-1)*(N-2))
plot(N,pr1,ylim=c(.35,.75),pch=16,ylab="P(Y=1)")
```



4. Prob. 3.122.- Llegan clientes a un mostrador de salida en una tienda de departamentos de acuerdo con una distribución de Poisson, a un promedio de siete por hora. Durante una hora determinada, ¿cuáles son las probabilidades de que
   (a) no lleguen más de tres clientes?,
   (b) lleguen al menos dos clientes?,
   (c) lleguen exactamente cinco clientes?

   Sea $Y =$ número de clientes llegando al mostrador en una hora.

$$Y \sim \text{Poisson}(\lambda = 7)$$

(a)

$$P(Y \leq 3) = \sum_{k=0}^{3} \frac{\lambda^k \, e^{-\lambda}}{k!} = e^{-7}\left(\frac{7^0}{0!} + \frac{7^1}{1!} + \frac{7^2}{2!} + \frac{7^3}{3!}\right) = e^{-7}\left(1 + 7 + \frac{49}{2} + \frac{343}{6}\right)$$

```
exp(-7)*(1+7+49/2+343/6) # = 0.08176542
ppois(q = 3, lambda = 7) # = 0.08176542
```

(b)

$$P(Y \geq 2) = 1 - P(Y \leq 1) = 1 - \sum_{k=0}^{1} \frac{\lambda^k \, e^{-\lambda}}{k!} = 1 - e^{-7}\left(\frac{7^0}{0!} + \frac{7^1}{1!}\right) = 1 - e^{-7}\,(1 + 7)$$

```
1 - exp(-7)*(1+7) # = 0.9927049
1 - ppois(q = 1, lambda = 7) # = 0.9927049
```

(c)

$$P(Y = 5) = \frac{\lambda^5 \, e^{-\lambda}}{5!}$$

```
exp(-7)*(7^5)/factorial(5) # = 0.1277167
dpois(x = 5, lambda = 7)   # = 0.1277167
```

5. Prob. 3.125.- Consulte el Ejercicio 3.122. Si se requieren alrededor de diez minutos para servir a cada cliente, encuentre la media y la varianza del tiempo total de servicio para clientes que lleguen durante un periodo de 1 hora. (Suponga que hay un número suficiente de dependientes para que el cliente no tenga que esperar ser atendido.) ¿Es probable que el tiempo total de servicio exceda de 2.5 horas?

   Si $Y =$ número de clientes llegando al mostrador en una hora, entonces $T = 10 * Y$ sería el tiempo total de servicio para los clientes que llegan durante un periodo de una hora. La media y varianza de T son

$$\mathsf{E}(T) = \mathsf{E}(10\,Y) = 10\,\mathsf{E}(Y) = 10 \times 7 = 70\,\text{min}$$
$$\mathsf{Var}(T) = \mathsf{Var}(10\,Y) = 100\,\mathsf{Var}(Y) = 100 \times 7 = 700\,\text{min}$$

   Por otro lado,

$$P(T > 2.5\,\text{hrs}) = P(T > 150) = P(Y > 150/10) = P(Y > 15) = 1 - P(Y \leq 15) = 0.0024$$

   entonces, no es probable que se tenga un tiempo total de servicio de más de 2.5 horas.

6. Prob. 3.126.- Consulte el Ejercicio 3.122. Suponga que ocurren llegadas de acuerdo con un proceso de Poisson con un promedio de siete por hora. ¿Cuál es la probabilidad de que exactamente dos clientes lleguen en dos horas entre

   (a) las 2:00 p.m. y las 4:00 p.m. (un periodo continuo de dos horas)?,

   (b) la 1:00 p.m. y las 2:00 p.m. o entre las 3:00 p.m. y las 4:00 p.m. (dos periodos de una hora separados que totalizan dos horas)?

   La respuesta debe ser la misma en ambos casos.

   (a) Sea $Y =$ número de clientes que llegan en un periodo de dos horas. Entonces

$$Y \sim \text{Poisson}(\lambda = 14)$$

   Entonces

$$P(Y = 5) = \frac{\lambda^5\, e^{-\lambda}}{5!} = \frac{14^5\, e^{-14}}{5!}$$

```
exp(-14)*(14^5)/factorial(5) # = 0.003726801
dpois(5,14)                  # = 0.003726801
```

   (b) Sea $Y = Y_1 + Y_2$, donde $Y_1 =$ número de clientes que llegan en el primer periodo, y $Y_2 =$ el número que llega en el segundo. Suponemos Y_1 y Y_2 independientes. Entonces

$$Y_1 \sim \text{Poisson}(\lambda_1 = 7) \quad \text{y} \quad Y_2 \sim \text{Poisson}(\lambda_2 = 7)$$

   De aquí que

$$\begin{aligned} P(Y = 5) = &P(Y_1 = 0)P(Y_2 = 5) + P(Y_1 = 1)P(Y_2 = 4) + P(Y_1 = 2)P(Y_2 = 3) \\ &+ P(Y_1 = 3)P(Y_2 = 2) + P(Y_1 = 4)P(Y_2 = 1) + P(Y_1 = 5)P(Y_2 = 0) \end{aligned}$$

(¿Porqué?)

$$\begin{aligned}
P(Y=5) &= \frac{7^0\,e^{-7}}{0!}\,\frac{7^5\,e^{-7}}{5!} + \frac{7^1\,e^{-7}}{1!}\,\frac{7^4\,e^{-7}}{4!} + \frac{7^2\,e^{-7}}{2!}\,\frac{7^3\,e^{-7}}{3!} \\
&\quad + \frac{7^3\,e^{-7}}{3!}\,\frac{7^2\,e^{-7}}{2!} + \frac{7^4\,e^{-7}}{4!}\,\frac{7^1\,e^{-7}}{1!} + \frac{7^5\,e^{-7}}{5!}\,\frac{7^0\,e^{-7}}{0!} \\
&= 7^5\,e^{-14}\left(\frac{1}{0!\,5!} + \frac{1}{1!\,4!} + \frac{1}{2!\,3!} + \frac{1}{3!\,2!} + \frac{1}{4!\,1!} + \frac{1}{5!\,0!}\right) \\
&= \frac{7^5\,e^{-14}}{5!}\left(\frac{5!}{0!\,5!} + \frac{5!}{1!\,4!} + \frac{5!}{2!\,3!} + \frac{5!}{3!\,2!} + \frac{5!}{4!\,1!} + \frac{5!}{5!\,0!}\right) \\
&= \frac{7^5\,e^{-14}}{5!}\left(\binom{5}{0} + \binom{5}{1} + \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5}\right) \\
&= \frac{7^5\,e^{-14}}{5!}\,2^5 = \frac{14^5\,e^{-14}}{5!}
\end{aligned}$$

Esto es lo mismo que en el inciso anterior. En realidad, esta es una propiedad general: La suma de k Poissones independientes, con parámetros $\lambda_1, \ldots, \lambda_k$ respectivamente, es una variable Poisson de parámetro $\lambda_1 + \cdots + \lambda_k$. En particular, si $Y_1 \sim$ Poisson(λ) y $Y_2 \sim$ Poisson(λ), entonces $Y = Y_1 + Y_2 \sim$ Poisson(2λ), asumiendo independencia entre Y_1 y Y_2.

7. Prob. 3.128.- Llegan autos a una caseta de pago de peaje de acuerdo con un proceso de Poisson con media de 80 autos por hora. Si el empleado hace una llamada telefónica de 1 minuto, ¿cuál es la probabilidad de que al menos 1 auto llegue durante la llamada?

   Sea Y = número de autos que llegan a la caseta en 1 minuto. Entonces $Y \sim$ Poisson$(80/60)$. De aquí que

   $$P(Y \geq 1) = 1 - P(Y=0) = 1 - \frac{e^{-4/3}(4/3)^0}{0!} = 1 - e^{-4/3} = 0.7364$$

8. Prob. 3.129.- Consulte el Ejercicio 3.128. ¿Cuánto puede durar la llamada telefónica del empleado si la probabilidad es al menos .4 de que no lleguen autos durante la llamada?

   Sea d = la duración de la llamada (en minutos), y sea Y = número de llamadas que llegan en esos d minutos. Entonces $Y \sim$ Poisson($d \times 80/60$). Queremos que $P(Y=0) \geq 0.4$:

   $$0.4 \leq P(Y=0) = e^{-4d/3} \quad \Rightarrow \quad -\frac{4}{3}d \geq \log(0.4) \quad \Rightarrow \quad d \leq -\frac{3}{4}\log(0.4) = 0.687218$$

   esto es, la llamada debe durar a lo más 41 segundos (la respuesta en el libro no es correcta).

9. Prob. 3.137.- La probabilidad de que un ratón inoculado con un suero contraiga cierta enfermedad es 0.2. Usando la aproximación de Poisson, encuentre la probabilidad de que al menos 3 de entre 30 ratones inoculados contraigan la enfermedad.

   Sea Y = número de ratones que contraen la enfermedad. Entonces, es razonable que

   $$Y \sim \text{Binomial}(n=30, p=0.2)$$

   de aquí que

   $$\begin{aligned}
   P(Y \geq 3) = 1 - P(Y \leq 2) &= 1 - \left(\binom{30}{0} 0.2^0\, 0.8^{30} + \binom{30}{1} 0.2^1\, 0.8^{29} + \binom{30}{2} 0.2^2\, 0.8^{28}\right) \\
   &= 0.9558
   \end{aligned}$$

```
p = 0.2
1-(choose(30,0)*p^0*(1-p)^30+choose(30,1)*p^1*(1-p)^29+choose(30,2)*p^2*(1-p)^28 )
# 0.955821
1 - pbinom(q=2, size=30, prob=p)  # = 0.955821
```

(la respuesta en el libro se refiere al caso: A lo más 3 enfermos: `ppois(3,6) = 0.1512039`).

Ahora, si usamos la aproximación Poisson,

$$Y \dot{\sim} \text{Poisson}(\lambda = np) = \text{Poisson}(\lambda = 6)$$

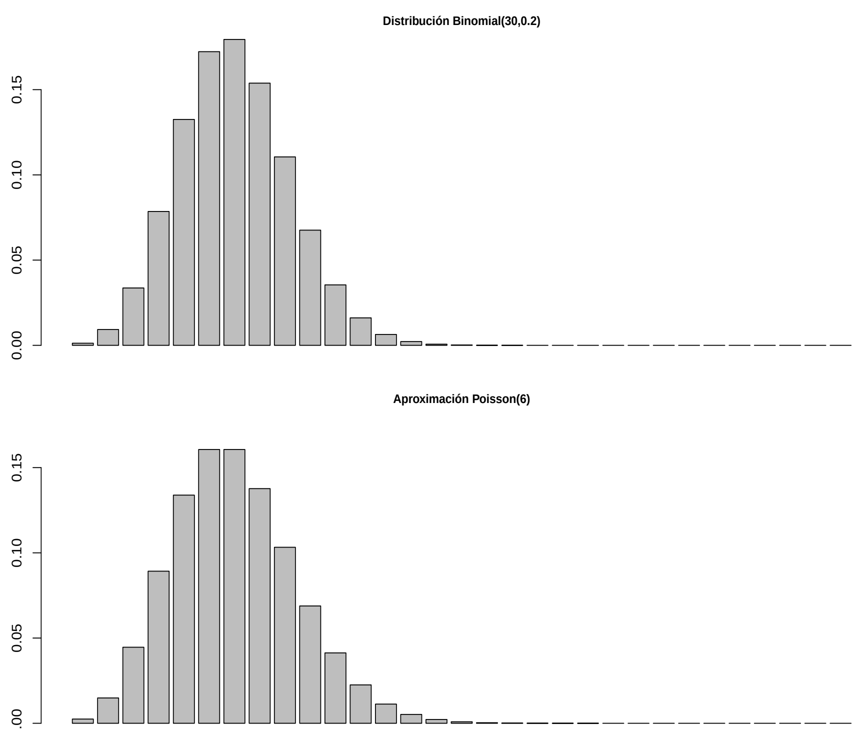
entonces

$$P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - \left(\frac{e^{-6}\, 6^0}{0!} + \frac{e^{-6}\, 6^1}{1!} + \frac{e^{-6}\, 6^2}{2!} \right) = 1 - e^{-6}\, (1 + 6 + 18) = 0.938$$

```
1-exp(-6)*(1+6+18) # 0.9380312
1-ppois(2,6)       # 0.9380312
```

Una comparación a más detalle entre ambas distribuciones

```
# Comparación Binomial(n,p) con Poisson(np)
n   = 30
p   = 0.2
lam = n*p
x   = 0:n
pb  = dbinom(x,n,p)
pp  = dpois(x,lam)
par(mfcol=c(2,1),mar=c(2,2,2,2))
barplot(pb,ylim=c(0,.18),main="Distribución Binomial(30,0.2)",cex.main=.8)
barplot(pp,ylim=c(0,.18),main="Aproximación Poisson(6)",cex.main=.8)
```



10. Prob. 3.141.- Un fabricante de alimentos usa una máquina de moldeo por inyección (que produce galletas del tamaño de un bocado y botanas) que proporciona un ingreso para la empresa a razón de $200 por hora cuando está en operación. No obstante, la máquina se descompone a un promedio de dos veces por cada día que trabaja. Si Y denota el número de descomposturas por día, el ingreso diario generado por la máquina es $R = 1600 - 50Y^2$. Encuentre el ingreso diario esperado por usar la máquina.

$$\mathsf{E}(R) = \mathsf{E}(1600 - 50Y^2) = 1600 - 50\mathsf{E}(Y^2) = 1600 - 50\left(\,\mathsf{Var}(Y) + \mathsf{E}^2(Y)\,\right) = 1600 - 50(2+4) = 1300$$

11. Prob. 4.66.- Una operación de maquinado produce cojinetes con diámetros que están normalmente distribuidos con media de 3.0005 pulgadas y desviación estándar de .0010 pulgadas. Las especificaciones requieren que los diámetros de los cojinetes se encuentren en el intervalo $3.000 \pm .0020$ pulgadas. Los cojinetes que estén fuera de este intervalo son considerados de desecho y deben volver a maquinarse. Con el ajuste de la máquina existente, ¿qué fracción de la producción total se desechará?

    Sea $Y =$ diámetro de un cojinete genérico. La probabilidad de que una pieza se encuentre dentro de las especificaciones es:

$$\begin{aligned} P(3-0.002 \le Y \le 3+0.002) &= P(\ (3-0.002-3.0005)/.001 \le Z \le (3+0.002-3.0005)/.001) \\ &= P(\ -0.0025/.001 \le Z \le 0.0015/.001) \\ &= P(\ -2.5 \le Z \le 1.5) = 0.927 \end{aligned}$$

    de aquí que el 7.3% de los cojinetes se desecharán por estar fuera de especificaciones.

```
pnorm(1.5)-pnorm(-2.5)       # 0.9269831
1 - (pnorm(1.5)-pnorm(-2.5)) # 0.07301687
```

12. Prob. 4.75.- Una máquina expendedora de bebidas gaseosas puede ser regulada para descargar un promedio de μ onzas por vaso. Si las onzas están normalmente distribuidas con desviación estándar de 0.3 onzas, determine los valores para μ de modo que vasos de 8 onzas se sirvan (se derramen?) sólo 1% del tiempo.

    Sea $Y =$ cantidad servida. Queremos μ tal que $Y \sim N(\mu, \sigma = 0.3)$ y $P(Y > 8) = 0.01$. Sea Z_α tal que $P(Z < Z_\alpha) = 0.99$, entonces $P(Y < \mu +$

$$0.01 = P(Y > 8) = P\left(\frac{Y-\mu}{\sigma} > \frac{8-\mu}{\sigma}\right) \quad \Rightarrow \quad \frac{8-\mu}{\sigma} = Z_\alpha$$

    de aquí

$$\mu = 8 - \sigma\, Z_\alpha = 8 - 0.3\ (2.326348) = 7.302096$$

```
8 - 0.3 * qnorm(.99)  # 7.302096
```

13. Prob. 4.77.- Los exámenes de admisión SAT y ACT (de aptitud y universitario) se aplican a miles de estudiantes cada año. Las secciones de matemáticas de cada uno de estos exámenes producen califi caciones que están normalmente distribuidas, en forma aproximada. En años recientes las calificaciones de exámenes SAT de matemáticas han promediado 480 con desviación estándar de 100. El promedio y desviación estándar para calificaciones ACT de matemáticas son 18 y 6, respectivamente.

    (a) Una escuela de ingeniería establece 550 como calificación mínima SAT de matemáticas para estudiantes de nuevo ingreso. ¿Qué porcentaje de estudiantes obtendrá una calificación por debajo de 550 en un año típico?

(b) ¿Qué calificación debe establecer la escuela de ingeniería como estándar comparable en el examen ACT de matemáticas?

(a) $P(SAT < 550) = 0.758$

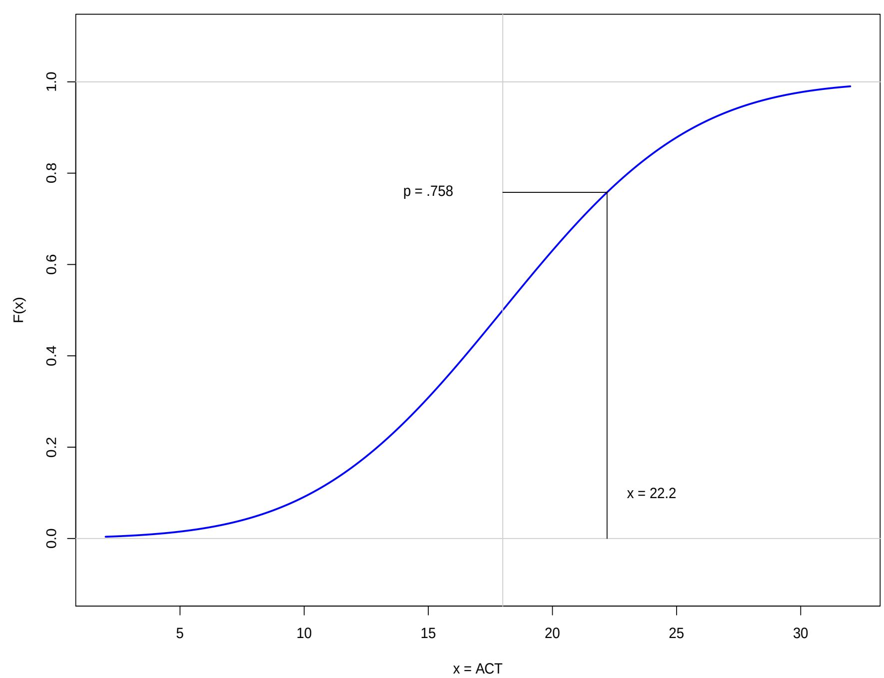
```
pnorm(q=550, mean=480, sd=100) # 0.7580363
pnorm((550-480)/100)           # 0.7580363
```

(b) Queremos a tal que $P(ACT < a) = 0.7580363$. Puede verse que el valor apropiado para a es 22.2

```
qnorm(p=0.7580363, mean=18, sd=6) # 22.2
18 + 6*qnorm(p=0.7580363)         # 22.2
```

Una gráfica de la función de distribución de la variable ACT:

```
M  = 201
x  = seq(2,32,length=M)
ff = pnorm(x,mean=18,sd=6)
plot(x,ff,type="l",col="blue",lwd=2,ylim=c(-0.1,1.1),xlab="x = ACT",ylab="F(x)")
abline(h=c(0,1),v=18,col=gray(.8))
segments(x0=18,y0=.7580363,x1=22.2,y1=.7580363)
segments(x0=22.2,y0=0,x1=22.2,y1=.7580363)
text(15,.758,"p = .758")
text(24,.1,"x = 22.2")
```



14. Prob. 4.80.- Suponga que Y está normalmente distribuida con media μ y desviación estándar σ. Después de observar el valor de Y, un matemático construye un rectángulo con longitud $L = |Y|$ y ancho $W = 3|Y|$. Denote con A el área del triángulo resultante. ¿Cuál es $\mathsf{E}(A)$?

$$\mathsf{E}(A) = \mathsf{E}(L\,W) = \mathsf{E}(3Y^2) = 3\mathsf{E}(Y^2) = 3\left(\mathsf{Var}(Y) + \mathsf{E}^2(Y)\right) = 3(\sigma^2 + \mu^2)$$