¿Determina el género el éxito en solicitudes para pruebas con el telescopio Hubble?

El Instituto de Ciencias del Telescopio Espacial recibe cada año más de 1000 solicitudes y tan solo 200 son aceptadas.

Según un estudio publicado por Neill Reid en 2014, las propuestas de investigación para utilizar el Hubble realizadas por hombres tuvieron una mayor aceptación que las de mujeres. Nuevos estudios, sugieren lo mismo para el Webb y otros ámbitos.

En este ejercicio se utilizan datos de 11 ciclos y si la proporción de propuestas aceptadas es significativamente mayor en un género que en otro.

Por tanto, se plantea la hipótesis contraria y se fija un nivel de aceptación de $\alpha = 0.05$ y un intervalo de confianza del 95%.

In [9]:
```
ciclos = c(11:20, 21)

pi_male = c(873, 840, 728, 580, 591, 650, 770, 824, 768, 832, 805); n1 = sum(pi_ma
pi_female = c(205, 205, 177, 145, 146, 171, 188, 226, 239, 253, 289); n2 = sum(pi_
tot = c(n1, n2); #tot

ac_male = c(168, 212, 171, 168, 171, 159, 188, 170, 160, 183, 189)
ac_female = c(30,29,39,39,32,32,40,26,39,48,63)

ex_ml = ac_male / pi_male; #ex_ml
ex_fm = ac_female / pi_female; #ex_fm

#n1; n2;
#length(yr); length(pi_male); length(pi_female); length(ac_male); length(ac_female
```

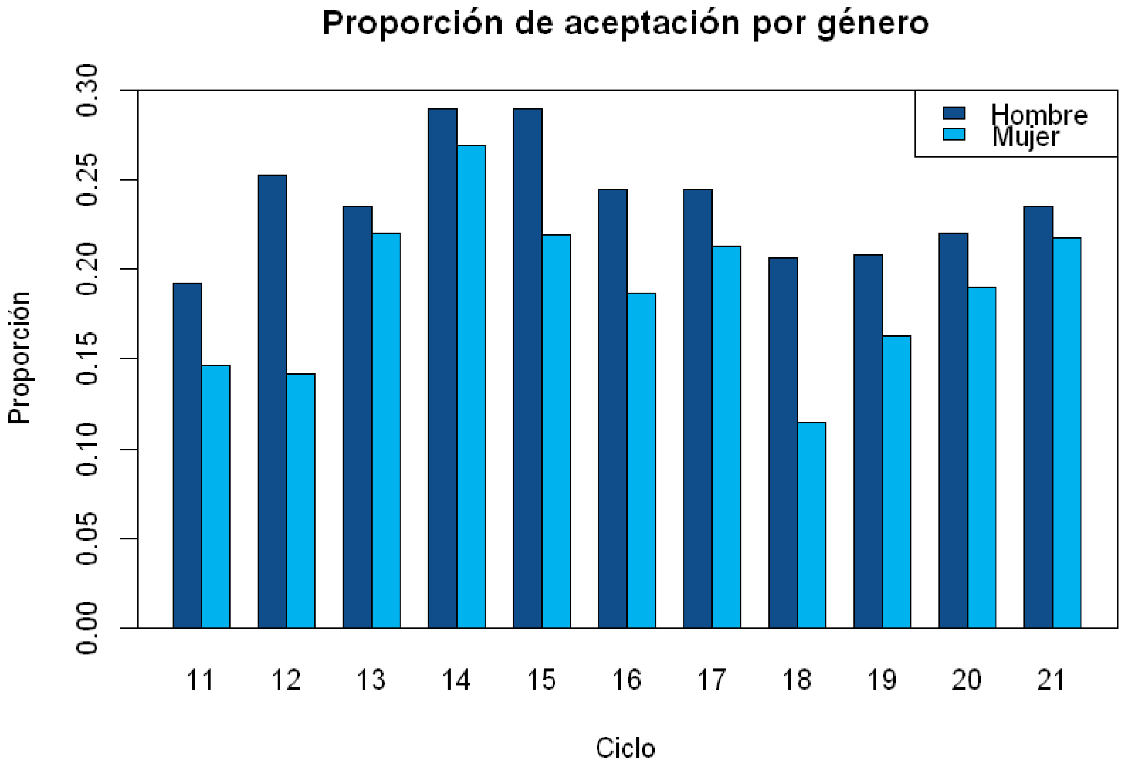
In [10]:
```
ex_ml = ac_male / pi_male; #ex_ml
ex_fm = ac_female / pi_female; #ex_fm

x = rbind(ex_ml, ex_fm); #x

options(repr.plot.height = 5, repr.plot.width = 7) #Se ajusta el tamaño

barplot(x, names.arg = ciclos, main = "Proporción de aceptación por género", ylim 
       xlab = "Ciclo", ylab = "Proporción", col = c("#104E8B", "#00B2EE"))

legend("topright", legend = c("Hombre", "Mujer"),
       fill = c("#104E8B", "#00B2EE"))
box()
```



La primera impresión a partir del gráfico, es que hay una diferencia en favor de un género. Veamos si pueden ser debidas al muestreo o no.

In [11]:
```
medx = c(mean(ex_ml),mean(ex_fm))
sdx = c(sd(ex_ml), sd(ex_fm))
acep = c(sum(ac_male), sum(ac_female))

df = data.frame(medx, sdx, c(n1, n2), acep, row.names = c("Hombre", "Mujer"))
names(df) = c("Media", "Desviación", "Peticiones", "Aceptadas")
df
```

	Media	Desviación	Peticiones	Aceptadas
Hombre	0.2378960	0.03146144	8261	1939
Mujer	0.1892844	0.04462227	2244	417

Intervalo de confianza

Como consecuencia del teorema del límite central, para cualesquiera distribuciones de partida la distribución muestral de la diferencia de medias puede aproximarse por una normal siempre que el tamaño de las muestras sea suficientemente grande.

Un caso particular de este resultado es el intervalo de confianza para la diferencia de proporciones de dos distribuciones binomiales (éxito/fracaso como son aceptada/rechazada).

Para el cálculo del intervalo de confianza, que fijamos en un 95%, se utiliza:

$$I = \left[\bar{P}_1 - \bar{P}_2 \pm z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}\right]$$

In [12]:
```
nc = 95
p1 = acep[1] / n1; p2 = acep[2] / n2

zlim = (1 - nc/100)/2; #zlim = 0.05/2 que sería igual a alpha = 0.05
zint = qnorm(zlim, lower.tail=FALSE)

l = zint * sqrt( (p1 * (1 - p1)/n1) + (p2 * (1 - p2)/n2) )
l1 = p1 - p2 - l
l2 = p1 - p2 + l

cat("El intervalo de confianza del", nc, "% para la diferencia en proporción de ["
```

```
El intervalo de confianza del 95 % para la diferencia en proporción de [ 0.0303809
3 , 0.06739601 ]
```

El hecho de que un intervalo de confianza no contenga el cero significa que se puede afirmar con un nivel de confianza del 95% que la diferencia significativa entre las dos proporciones que se están comparando, **no son debidas al azar**.

Contraste de hipótesis

Al ser el número de datos de la muestra grande, se aproxima la distribución binomial por una normal.

El contraste de igualdad entre las dos propociones, con p_1 la proporción de éxito entre hombres y p_2 entre mujeres, se plantea con la hipótesis nula como lo contrario a lo que dicen los datos:

$$H_0 = p_1 \leq p_2$$

$$H_0 = p_1 > p_2$$

El contraste es unliteral y el nivel de aceptaión fijado de $\alpha = 0.05$

In [13]:
```
alpha = 0.05
p1 = acep[1] / n1; p2 = acep[2] / n2
z.alpha = qnorm(alpha, lower.tail = FALSE)

z = (p1 - p2) / (sqrt(p1*(1-p1)/(n1) + p2*(1-p2)/(n2))) #Estadístico

z < z.alpha

#El nivel para el que se rechaza al hipótesis:
pnorm(z, lower.tail = FALSE)
```

FALSE

1.1254270313962e-07

Por tanto, se rechaza la hipótesis nula para un nivel del significación de 0.05. La proporción de solicitudes aceptadas en hombres no es significativamente menor que en mujeres, o equivalentemente, tienen más proporción de éxito.

Además, el nivel para el que se rechaza la hipótesis nula es muy pequeño $\sim 10^{-7}$.

R dispone de una función propia para calcular intervalos de confianza y contrastar hipótesis. En este caso, vamos a suponer por hipótesis nula, que las proporciones son iguales.

In [14]:
```
#prop.test(acep, tot, alternative= "greater", conf.level = 0.95)
prop.test(x=c(1939,417),n=c(8261,2244),conf.level=0.95, alternative="two.sided")
```

```
	2-sample test for equality of proportions with continuity correction

data:  c(1939, 417) out of c(8261, 2244)
X-squared = 23.963, df = 1, p-value = 9.822e-07
alternative hypothesis: two.sided
95 percent confidence interval:
 0.03009758 0.06767936
sample estimates:
   prop 1    prop 2 
0.2347173 0.1858289 
```

El p-value cambia ligeramente (por el uso de X-squared) pero el orden $\sim 10^{-7}$ se mantiene, como nuestro nivel de significación es mayor, nos lleva a rechazar la hipótesis nula y a "aceptar" que la proporción de éxito no es significativamente igual en ambos géneros para un alpha prefijado de 0.05. Por otro lado, el intervalo de confianza no incluye el 0, las diferencias son debidas a un sesgo.

Una buena forma de hacerlo más justo es posible mediante solicitudes anónimas como se demuestra en [2].

Otro ejercicio que podría ser interesante, sería añadir en la comparativa las peticiones por países, continentes o edades.

Digitalización de una gráfica.

Para una gráfica encontrada posterior, con datos más antiguos, se ha hecho uso del software *https://apps.automeris.io/wpd/* para extraer los datos de https://pubs.aip.org/physicstoday/Online/30350/Doling-out-Hubble-time-with-dual-anonymous, donde en 2018 se implementó el sistema anónimo y las mujeres obtuvieron mejor resultado que los hombres.

In [15]:
```
data = read.table(file = "DatosHubble.txt",   # Archivo de datos TXT/.dat indicado
           header = TRUE,       # Si se muestra el encabezado (TRUE) o no (FALSE)
           #sep = " ",           # Separador de las columnas del archivo
           dec = ",")           # Separador decimal (en DatosHubble.txt se guardó 

names(data) = c("Hombre", "Mujer", "Año")
head(data)
```

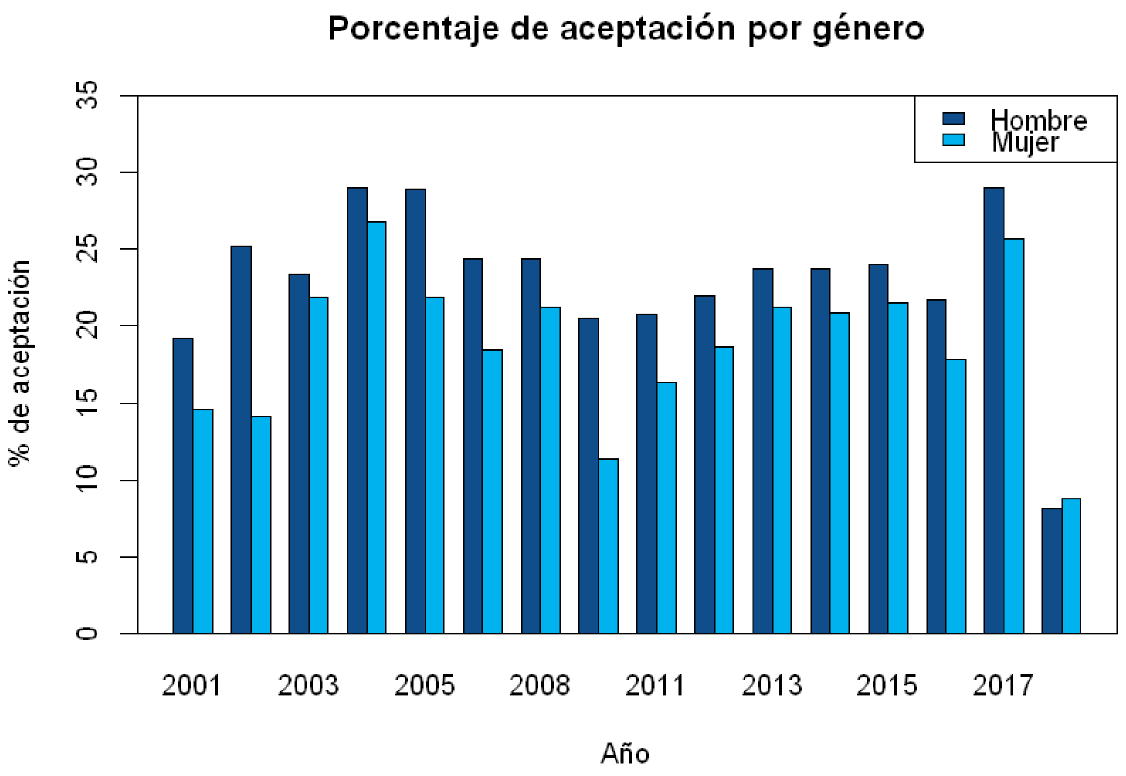
Hombre	Mujer	Año
19.20771	14.58244	2001
25.18201	14.13276	2002
23.38330	21.90578	2003
28.97216	26.78801	2004
28.90792	21.90578	2005
24.41113	18.50107	2006

In [16]:
```
male = data$Hombre; female = data$Mujer
x_new = rbind(male, female); #x_new
an = data$Año; #an; length(an)

options(repr.plot.height = 5, repr.plot.width = 7) #Se ajusta el tamaño

barplot(x_new, names.arg = an, main = "Porcentaje de aceptación por género", ylim =
       xlab = "Año", ylab = "% de aceptación", col = c("#104E8B", "#00B2EE"))

legend("topright", legend = c("Hombre", "Mujer"),
       fill = c("#104E8B", "#00B2EE"))
box()
```



[BIBLIOGRAFÍA]

[1] Neill Reid. 2014, Gender-based Systematics in HST Proposal Selection.

[2] Gareth Hunt, Frederic R. Schwab and Patricia A. Henning NRAO. 2021, Gender-Related Systematics in the NRAO Proposal Review Process.

[3] Stefanie K. Johnson1 and Jessica F. Kirk2. 2020, Dual-anonymization Yields Promising Results for Reducing Gender Bias: A Naturalistic Field Experiment of Applications for Hubble Space Telescope Time.

[ANEXO]

El último estudio encontrado añade 2019 y saca mejor proporción los hombres, con selección anónima habría que esperar algunos años para un mejor análisis. http://surveygizmoresponseuploads.s3.amazonaws.com/fileuploads/623127/5043187/180-69eaba1161478c402d5ffa2138dc3668_StrolgerLou.pdf