Zadanie 2. W tym zadaniu zajmować będziemy się wyznaczaniem przedziału ufności dla różnicy dwóch średnich w modelu normalnym o **znanych wariancjach** σ_1^2, σ_2^2 na poziomie ufności $0.95 =$ 1 - α. W tym celu korzystać będziemy z następującego wzoru

$$\overline{x_1} - \overline{x_2} \pm z_{1-\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

```
known_var <- function(x1, x2, var1, var2) return(c(mean(x1) - mean(x2)
                - qnorm(1 - .05/2) * sqrt(var1/length(x1) + var2/length(x2)),
                mean(x1) - mean(x2) + qnorm(1 - .05/2) * sqrt(var1/length(x1)
                + var2/length(x2))))
```

Zauważmy, że powyższy wzór nie ma zastosowania w przypadku rozkładów Cauchy'ego, ponieważ wariancje są dla nich nieokreślone.

Zaczniemy od różnych modeli pochodzących z rozkładu normalnego. Dla pojedynczyc prób 50-elementowych otrzymujemy następujące przedziału ufności

	$N(0,1), N(0,1)$	$N(0,1), N(1,1)$	$N(0,1), N(0,2)$	$N(0,1), N(1,2)$
dolna granica	-0.611	-1.439	-1.174	-1.616
górna granica	0.173	-0.655	0.066	-0.377

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	$N(0,1), N(0,1)$	$N(0,1), N(1,1)$	$N(0,1), N(0,2)$	$N(0,1), N(1,2)$
p	0.951	0.952	0.951	0.958

Możemy zauważyć, że w tym przypadku wraz ze wzrostem skali w drugim modelu wzrasta szerokość przedziału. Dodatkowo w każdym z przypadków prawdopodobieństwo pokrycia przez przedział różnicy średnich jest zbliżone do 0.95 t.j. wartości poziomu ufności.

Powtarzamy powyższy eksperyment dla modeli z rozkładu logistycznego. Otrzymujemy następujące przedziały ufności

	$L(0,1), L(0,1)$	$L(0,1), L(1,1)$	$L(0,1), L(0,2)$	$L(0,1), L(1,2)$
dolna granica	-0.284	-1.485	0.324	-2.621
górna granica	1.138	-0.064	2.573	-0.373

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	$L(0,1), L(0,1)$	$L(0,1), L(1,1)$	$L(0,1), L(0,2)$	$L(0,1), L(1,2)$
p	0.954	0.952	0.953	0.95

Podobnie jak wcześniej, wraz ze wzrostem skali w drugim modelu, wzrasta szerokość przedziału. Oczywiście wynika to wprost z postaci stosowanego przez nas wzoru. Dodatkowo prawopodobieństwo pokrycia nieznanej różnicy średnich przez przedział znów zbliżone jest do wartości poziomu ufności t.j. 0.95.

Zadanie 4. W tym zadaniu będziemy zajmować się wyznaczaniem przedziału ufności dla różnicy dwóch średnich w modelu normalnym o **nieznanych równych wariancjach** na poziomie ufności 0.95. W tym celu wykorzystywać będziemy następujący wzór

$$\overline{x_1} - \overline{x_2} \pm t_{1-\alpha/2, n_1+n_2-2} s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

gdzie

$$s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

```
unknown_eq_var <- function(x1, x2) {
  s <- sqrt(((length(x1) - 1)*sd(x1)^2 + (length(x2) - 1)*sd(x2)^2)/
  (length(x1) + length(x2) - 2))

  return(c(mean(x1) - mean(x2) - qt(1 - .05/2, length(x1) + length(x2) - 2)
  * sqrt(1/length(x1) + 1/length(x2))*s,
  mean(x1) - mean(x2) + qt(1 - .05/2, length(x1) + length(x2) - 2)
  * sqrt(1/length(x1) + 1/length(x2))*s))
}
```

W wykonywanym eksperymencie będziemy, analogicznie do zadania 2., dla wybranych rozkładów generować 50 obserwacji, które umożliwią nam wyznaczenie przedziału ufności dla różnicy średnich $\mu_1 - \mu_2$. Następnie będziemy powtarzać to doświadczenie 10 000 razy, aby móc oszacować prawdopodobieństwo pokrycia nieznanej średniej przez przedział ufności.

```
prob <- function(loc1, loc2, scale1, scale2, mean_dif, distr, n) {
  k <- 0
  for(i in c(1:10000)) {
    if(unknown_eq_var(distr(n, loc1, scale1), distr(n, loc2, scale2))[1] <=
    mean_dif &
       unknown_eq_var(distr(n, loc1, scale1), distr(n, loc2, scale2))[2] >=
    mean_dif) k <- k+1
  }
  return(k/10000)
}
```

Zaczniemy od rozkładów normalnych $N(0,1)$, $N(0,1)$. Otrzymany przedział ufności dla różnicy średnich wynosi

	$N(0,1)$
dolna granica	-0.424
górna granica	0.381

Powtarzamy powyższy eksperyment 10 000 razy i wyznaczamy prawdopodobieństwo pokrycia nieznanej średniej przez przedział ufności.

	$N(0,1)$
prawdopodobieństwo	0.952

Przejdziemy teraz do powtórzenia powyższego rozumowania dla rozkładów $L(0,1)$, $L(0,1)$. Otrzymujemy następujące wartości

	$L(0,1)$
dolna granica	-0.627
górna granica	0.822

	$L(0,1)$
prawdopodobieństwo	0.951

Widzimy więc, że w obu przypadkach prawdopodobieństwo pokrycia różnicy średnich jest bliskie 0.95 t.j. wartości poziomu ufności. Otrzymane przedziały są dość szerokie oraz przedział uzyskany dla rozkładu logistycznego jest szerszy od tego dla rozkładu normalnego.

Na koniec powtórzymy pierwszą część powyższych eksperymentów dla różnych modeli rozkładu Cauchy'ego, dla którego wariancja ani średnia nie jest zdefiniowana. Otrzymujemy następujące przedziały ufności:

	$C(0,1), C(0,1)$	$C(0,1), C(1,1)$	$C(0,1), C(0,2)$	$C(0,1), C(1,2)$
dolna granica	-4.560	-1.953	-2.429	-9.151
górna granica	2.136	1.919	3.917	7.153

Możemy zauważyć, że przedziały są bardzo szerokie i nie występuje pomiędzy nimi żadna zależność.

Zadanie 6. W tym zadaniu będziemy zajmować się wyznaczaniem przedziału ufności dla różnicy dwóch średnich w modelu normalnym o **nieznanych różnych wariancjach** na poziomie ufności 0.95. W tym celu wykorzystywać będziemy następujący wzór

$$\overline{x_1} - \overline{x_2} \pm t_{1-\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

gdzie

$$v = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^2}{n_2^2(n_2-1)}}$$

```
unknown_ineq_var <- function(x1, x2) {
  v <- (sd(x1)^2/length(x1) + sd(x2)^2/length(x2)^2)/
        (sd(x1)^4/(length(x1)^2 *(length(x1)-1)) + sd(x2)^4/(length(x2)^2
        *(length(x2)-1)))
  return(c(mean(x1) - mean(x2) - qt(1 - .05/2, v) *   sqrt(sd(x1)^2/length(x1)
         + sd(x2)^2/length(x2)), mean(x1) - mean(x2) + qt(1 - .05/2, v)
         * sqrt(sd(x1)^2/length(x1) + sd(x2)^2/length(x2))))
}
```

Zaczniemy od rozkładów normalnych $N(0,1)$, $N(0,1)$. Otrzymany przedział ufności dla różnicy średnich wynosi

	$N(0,1)$
dolna granica	-0.132
górna granica	0.674

Powtarzamy powyższy eksperyment 10 000 razy i wyznaczamy prawdopodobieństwo pokrycia nieznanej średniej przez przedział ufności.

	$N(0,1)$
prawdopodobieństwo	0.95

Przejdziemy teraz do powtórzenia powyższego rozumowania dla rozkładów $L(0,1)$, $L(0,1)$. Otrzymujemy następujące wartości

	$L(0,1)$
dolna granica	-0.822
górna granica	0.605

	$L(0,1)$
prawdopodobieństwo	0.948

Widzimy więc, że w obu przypadkach, podobnie jak w poprzednim zadaniu, prawdopodobieństwo pokrycia różnicy średnich jest bliskie 0.95 t.j. wartości poziomu ufności. Otrzymane przedziały znów są dość szerokie oraz przedział uzyskany dla rozkładu logistycznego jest szerszy od tego dla rozkładu normalnego.

Na koniec powtórzymy pierwszą część powyższych eksperymentów dla różnych modeli rozkładu Cauchy'ego, dla którego wariancja ani średnia nie są zdefiniowane. Otrzymujemy następujące przedziały ufności:

	$C(0,1), C(0,1)$	$C(0,1), C(1,1)$	$C(0,1), C(0,2)$	$C(0,1), C(1,2)$
dolna granica	-5.110929e+13	-6.360	-Inf	-123.666
górna granica	5.110929e+13	7.114	Inf	103.280

Możemy zauważyć, że pomiedzy otrzymanymi przedziałami nie występuje żadna zależność i są one niezwykle szerokie.

Zadanie 10. W tym zadaniu zajmować będziemy się wyznaczaniem przedziału ufności dla ilorazu dwóch wariancji w modelu normalnym o **nieznanych średnich** na poziomie ufności 0.95 = 1 - α. W tym celu korzystać będziemy z następującego wzoru

$$\left[\frac{s_1^2}{s_2^2}\frac{1}{F_{\alpha/2,n_1-1,n_2-1}}, \frac{s_1^2}{s_2^2}F_{\alpha/2,n_1-1,n_2-1}\right]$$

```
unknown_mean <- function(x1, x2) return(c(sd(x1)^2* qf(.05/2, length(x1) - 1,
length(x2) - 1)/sd(x2)^2, sd(x1)^2/(sd(x2)^2 * qf(.05/2, length(x1) - 1,
length(x2) - 1))))
```

Zaczniemy od różnych modeli pochodzących z rozkładu normalnego. Dla pojedynczyc prób 50-elementowych otrzymujemy następujące przedziału ufności

	$N(0,1), N(0,1)$	$N(0,1), N(1,1)$	$N(0,1), N(0,2)$	$N(0,1), N(1,2)$
dolna granica	0.503	0.576	0.119	0.121
górna granica	1.563	1.790	0.368	0.377

Natomiast prawdopodobieństwo pokrycia nieznanej wartości ilorazu wynosi odpowiednio

	$N(0,1), N(0,1)$	$N(0,1), N(1,1)$	$N(0,1), N(0,2)$	$N(0,1), N(1,2)$
p	0.951	0.949	0.949	0.953

Możemy zauważyć, że wraz ze wzrostem skali w drugim modelu maleje szerokość przedziału. Dodatkowo w każdym z przypadków prawdopodobieństwo pokrycia przez przedział ilorazu wariancji jest zbliżone do 0.95 t.j. wartości poziomu ufności.

Powtarzamy powyższy eksperyment dla modeli z rozkładu logistycznego. Otrzymujemy następujące przedziały ufności

	$L(0,1), L(0,1)$	$L(0,1), L(1,1)$	$L(0,1), L(0,2)$	$L(0,1), L(1,2)$
dolna granica	0.688	0.583	0.161	0.053
górna granica	2.136	1.810	0.500	0.166

Natomiast prawdopodobieństwo pokrycia nieznanej wartości ilorazu wynosi odpowiednio

	$L(0,1), L(0,1)$	$L(0,1), L(1,1)$	$L(0,1), L(0,2)$	$L(0,1), L(1,2)$
p	0.886	0.889	0.893	0.886

Widzimy że, podobnie jak w przypadku rozkładu normalnego, wraz ze wzrostem skali w drugim modelu maleje szerokość przedziału. Dodatkowo dla omawianych rozkładów logistycznych prawdopodobieństwo pokrycia nieznanej wartości ilorazu wariancji przez przedział ufności jest mniejsze niż w przypadku rozkładów normalnych. W przybliżeniu wynosi jedynie 0.89.

Na koniec podamy wartości przedziałów skonstruowanych dla modeli z rozkładów Cauchy'ego. Nie będziemy jednak szacować prawdopodobieństwa pokrycia, ponieważ przypomnijmy, że dla rozkładu Cauchy'ego wartość wariancji jest nieokreślona.

	$C(0,1), C(0,1)$	$C(0,1), C(1,1)$	$C(0,1), C(0,2)$	$C(0,1), C(1,2)$
dolna granica	0.007	2.877	0.004	0.041
górna granica	0.022	8.933	0.014	0.127

Widzimy, że podobnie jak we wcześniejszych przykładach, dla rozkładów Cauchy'ego otrzymujemy szerokie i nieregularne przedziały ufności.

Zadanie 8. W tym zadaniu zajmować będziemy się wyznaczaniem przedziału ufności dla ilorazu dwóch wariancji w modelu normalnym o **znanych średnich** μ_1, μ_2 na poziomie ufności $0.95 = 1$ - α. W tym celu korzystać będziemy z następującego wzoru

$$\left[\frac{r_1}{r_2}\frac{1}{F_{\alpha/2,n_1-1,n_2-1}}, \frac{r_1}{r_2}F_{\alpha/2,n_1-1,n_2-1}\right],$$

gdzie

$$r_1 = \frac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{n_1-1} \text{ oraz } r_2 = \frac{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2}{n_2-1}.$$

```
var_modified <- function(x, mean) return(sum((x-mean)^2)/(length(x) - 1))

known_mean <- function(x1, x2, mean1, mean2) {
  return(c(var_modified(x1, mean1) * qf(.05/2, length(x1) - 1, length(x2) - 1)/
  var_modified(x2, mean2), var_modified(x1, mean1)/(var_modified(x2, mean2)
  * qf(.05/2, length(x1) - 1, length(x2) - 1))))
}
```

Zaczniemy od różnych modeli pochodzących z rozkładu normalnego. Dla pojedynczyc prób 50-elementowych otrzymujemy następujące przedziału ufności

	$N(0,1), N(0,1)$	$N(0,1), N(1,1)$	$N(0,1), N(0,2)$	$N(0,1), N(1,2)$
dolna granica	0.754	0.549	0.204	0.082
górna granica	2.343	1.704	0.635	0.256

Natomiast prawdopodobieństwo pokrycia nieznanej wartości ilorazu wynosi odpowiednio

	$N(0,1), N(0,1)$	$N(0,1), N(1,1)$	$N(0,1), N(0,2)$	$N(0,1), N(1,2)$
p	0.952	0.954	0.954	0.953

Możemy zauważyć, że i w tym przypadku wraz ze wzrostem skali w drugim modelu maleje szerokość przedziału. Dodatkowo w każdym z przypadków prawdopodobieństwo pokrycia przez przedział ilorazu wariancji jest zbliżone do 0.95 t.j. wartości poziomu ufności.

Powtarzamy powyższy eksperyment dla modeli z rozkładu logistycznego. Otrzymujemy następujące przedziały ufności

	$L(0,1), L(0,1)$	$L(0,1), L(1,1)$	$L(0,1), L(0,2)$	$L(0,1), L(1,2)$
dolna granica	1.190	0.636	0.236	0.186
górna granica	3.696	1.977	0.733	0.577

Natomiast prawdopodobieństwo pokrycia nieznanej wartości ilorazu wynosi odpowiednio

	$L(0,1), L(0,1)$	$L(0,1), L(1,1)$	$L(0,1), L(0,2)$	$L(0,1), L(1,2)$
p	0.894	0.89	0.893	0.894

Podobnie jak wcześniej, wraz ze wzrostem skali w drugim modelu, maleje szerokość przedziału. Widzimy jednak, że prawdopodobieństwo pokrycia dla każdej z rozważanych par rozkładów nie przekracza 0.9.

Zadanie 11. Przejdziemy do powtórzenia omawianych wcześniej eksperymentów dla prób 20- oraz 100-elementowych. Zaczniemy od eksperymentu z Zadania 2. Dla modeli z rozkładów $N(0,1)$, $N(0,1)$ otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	-0.526	-0.019
górna granica	0.714	0.535

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.948	0.949

Powtarzamy powyższy eksperyment dla modeli z rozkładów logistycznych $L(0,1), L(0,1)$. Otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	-1.86	-0.352
górna granica	-0.62	0.202

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.949	0.95

Podobnie jak wcześniej, prawopodobieństwo pokrycia nieznanej różnicy średnich przez przedział znów zbliżone jest do wartości poziomu ufności t.j. 0.95. Ponadto możemy zauważyć, że dla większej próby otrzymujemy węższe przedziały.

W dalszej części przejdziemy do eksperymentu z Zadania 4. Dla modeli z rozkładów $N(0,1)$, $N(0,1)$ otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	-0.944	-0.155
górna granica	0.398	0.383

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.953	0.951

Powtarzamy powyższy eksperyment dla z rozkładów logistycznych $L(0,1), L(0,1)$. Otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	-1.757	-0.427
górna granica	0.854	0.529

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.952	0.948

Widzimy, że znów wraz ze wzrostem liczebności próby zmniejsza się szerokość przedziału, zaś prawdopodobieństwo pokrycia nieznanej wartości różnicy przez przedział wynosi w przybliżeniu 0.95.

W następnym kroku powtórzymy eksperyment z Zadania 6. Dla modeli z rozkładów $N(0,1)$, $N(0,1)$ otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	-0.481	-0.155
górna granica	0.450	0.370

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.943	0.949

Powtarzamy powyższy eksperyment dla z rozkładów logistycznych $L(0,1), L(0,1)$. Otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	-1.676	-0.481
górna granica	0.565	0.462

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.952	0.951

Znowu możemy zauważyć, że wraz ze wzrostem liczebności próby zmniejsza się szerokość przedziału, zaś prawdopodobieństwo pokrycia nieznanej wartości różnicy przez przedział wynosi w przybliżeniu 0.95.

W następnym kroku powtórzymy eksperyment z Zadania 8. Dla modeli z rozkładów $N(0,1)$, $N(0,1)$ otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	0.492	0.693
górna granica	3.142	1.531

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.958	0.955

Powtarzamy powyższy eksperyment dla z rozkładów logistycznych $L(0,1), L(0,1)$. Otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	0.285	0.592
górna granica	1.819	1.308

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.906	0.885

Widzimy, że w kwestii szerokości przedziałów otrzymujemy wnioski analogiczne do poprzednio omawianych eksperymentów. Ponadto w przypadku rozkładu logistycznego prawdopodobieństwo pokrycia przez przedział jest mniejsze niż 0.95, wynosi w przybliżeniu 0.89. Dla rozkładu normalnego natomiast znów wynosi ono około 0.95.

Na koniec powatarzamy doświadczenie z Zadania 10. Dla modeli z rozkładów $N(0,1)$, $N(0,1)$ otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	0.331	0.446
górna granica	2.114	0.985

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.951	0.956

Powtarzamy powyższy eksperyment dla modeli z rozkładów logistycznych $L(0,1), L(0,1)$. Otrzymujemy następujące przedziały ufności

	20 obserwacji	100 obserwacji
dolna granica	0.296	0.787
górna granica	1.892	1.739

Natomiast prawdopodobieństwo pokrycia nieznanej wartości różnicy wynosi odpowiednio

	20 obserwacji	100 obserwacji
p	0.896	0.882

Widzimy, że w szerokość przedziałów znów maleje wraz ze wzrostem próby. Ponadto w przypadku rozkładu logistycznego prawdopodobieństwo pokrycia przez przedział jest mniejsze niż 0.95, wynosi w przybliżeniu 0.89. Dla rozkładu normalnego natomiast znów wynosi ono około 0.95.