Ejercicio 2

Analice los datos `Machines` del ejemplo guia 3 (ec. 4) a través de la función `lme`.

Modelo

un modelo para la k-ésima observación en el nivel i de efecto fijo A y el nivel j de efecto aleatorio B es

$y_{ijk} = \mu + \alpha_i + b_j + (ab)_{ij} + \epsilon_{ijk}$, (4)

donde:

$b_j \sim N(0, \sigma_b^2)$,

$(\alpha b)_{ij} \sim N(0, \sigma_{\alpha\beta}^2)$,

$\epsilon_{ijk} \sim N(0, \sigma^2)$,

y todas las v.a. son mutuamente independientes.

Ademas, μ es la media poblacional global, α_i son los I efectos fijos para el factor A, y b_j representan los J efectos aleatorios para el factor B.

Por otro lado, $(\alpha b)_i j$ son las IJ interacciones.

Datos

- El marco de datos Machines, del paquete `nlme`, contiene datos de un experimento industrial que compara 3 tipos de máquinas diferentes.
- El objetivo del experimento es determinar qué tipo de máquina daba como resultado la mayor productividad de los trabajadores.
- Se seleccionaron al azar 6 trabajadores para participar en la prueba, y cada trabajador operó cada máquina 3 veces (presumiblemente después de un período apropiado de capacitación diseñado para eliminar cualquier "efecto de aprendizaje").

(a) Intente encontrar el modelo más apropiado, teniendo cuidado de examinar los gráficos de verificación del modelo apropiado.

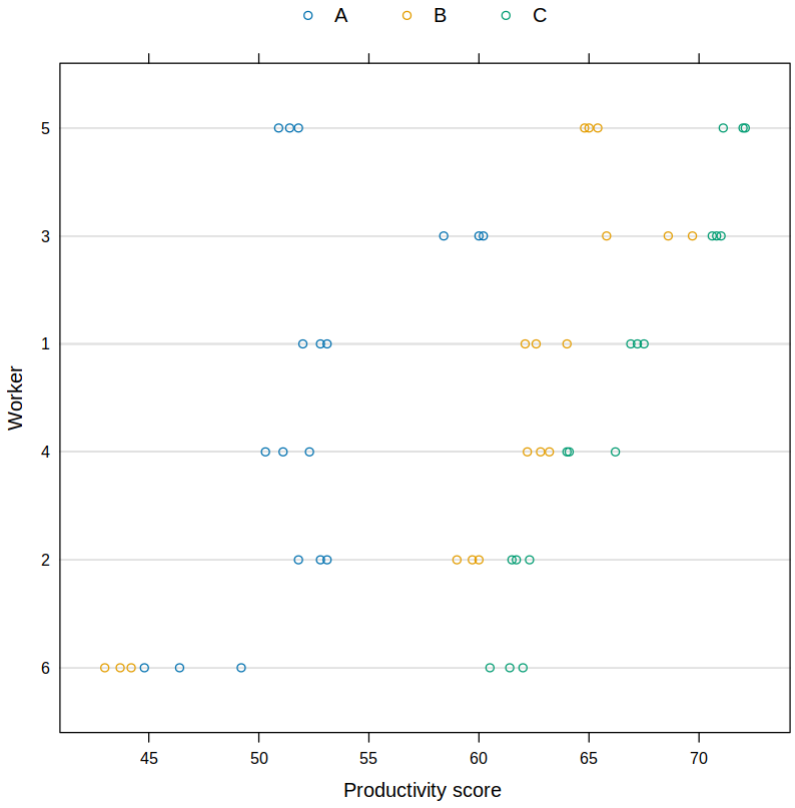
In [1]:
```
library(nlme)
attach(Machines)
Machines
```

A nffGroupedData: 54 × 3

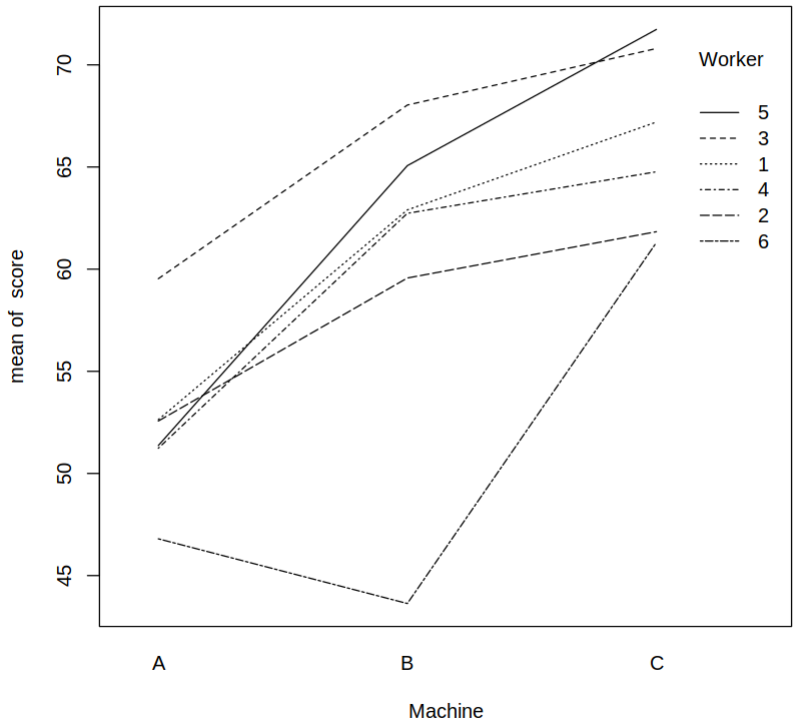
	Worker	Machine	score
	<ord>	<fct>	<dbl>
1	1	A	52.0
2	1	A	52.8
3	1	A	53.1
4	2	A	51.8
5	2	A	52.8
6	2	A	53.1
7	3	A	60.0
8	3	A	60.2
9	3	A	58.4
10	4	A	51.1
11	4	A	52.3
12	4	A	50.3
13	5	A	50.9
14	5	A	51.8
15	5	A	51.4
16	6	A	46.4
17	6	A	44.8
18	6	A	49.2
19	1	B	62.1
20	1	B	62.6
21	1	B	64.0
22	2	B	59.7
23	2	B	60.0
24	2	B	59.0
25	3	B	68.6
26	3	B	65.8
27	3	B	69.7
28	4	B	63.2
29	4	B	62.8
30	4	B	62.2
31	5	B	64.8
32	5	B	65.0

	Worker	Machine	score
	<ord>	<fct>	<dbl>
33	5	B	65.4
34	6	B	43.7
35	6	B	44.2
36	6	B	43.0
37	1	C	67.5
38	1	C	67.2
39	1	C	66.9
40	2	C	61.5
41	2	C	61.7
42	2	C	62.3
43	3	C	70.8
44	3	C	70.6
45	3	C	71.0
46	4	C	64.1
47	4	C	66.2
48	4	C	64.0
49	5	C	72.1
50	5	C	72.0
51	5	C	71.1
52	6	C	62.0
53	6	C	61.4
54	6	C	60.5

In [2]:
```
plot(Machines)
```



In [3]:
```
interaction.plot( Machine, Worker, score)
```

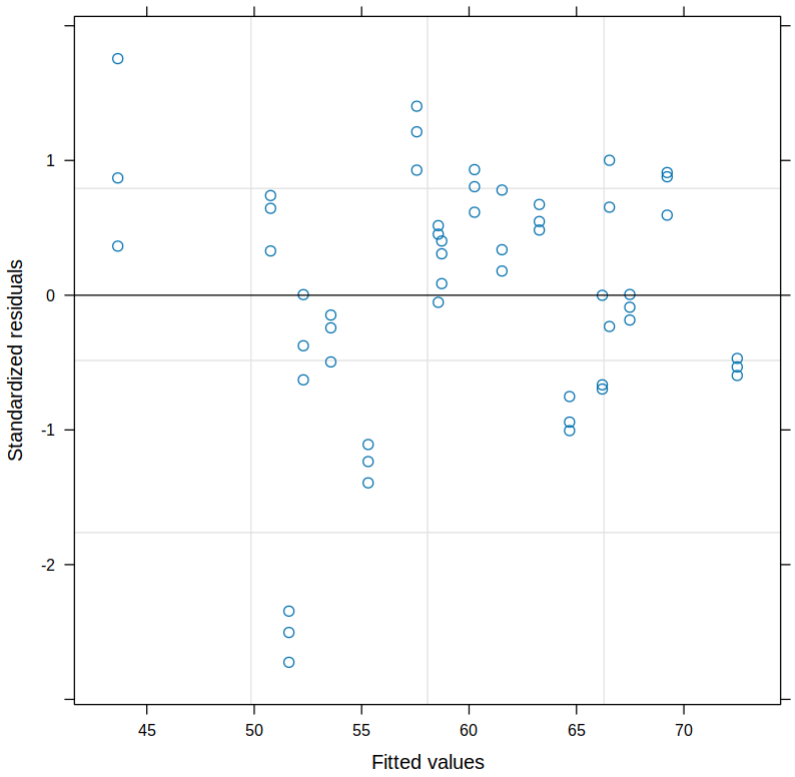


Construimos un primer modelo sin interaccion

In [4]:
```
m0 <- lme( score ~ Machine, data = Machines, random = ~ 1 | Worker )
m0
```

```
Linear mixed-effects model fit by REML
  Data: Machines
  Log-restricted-likelihood: -143.4391
  Fixed: score ~ Machine
(Intercept)    MachineB    MachineC
  52.355556    7.966667   13.916667

Random effects:
 Formula: ~1 | Worker
        (Intercept) Residual
StdDev:    5.146552 3.161647

Number of Observations: 54
Number of Groups: 6
```

In [5]:
```
plot(m0)
```

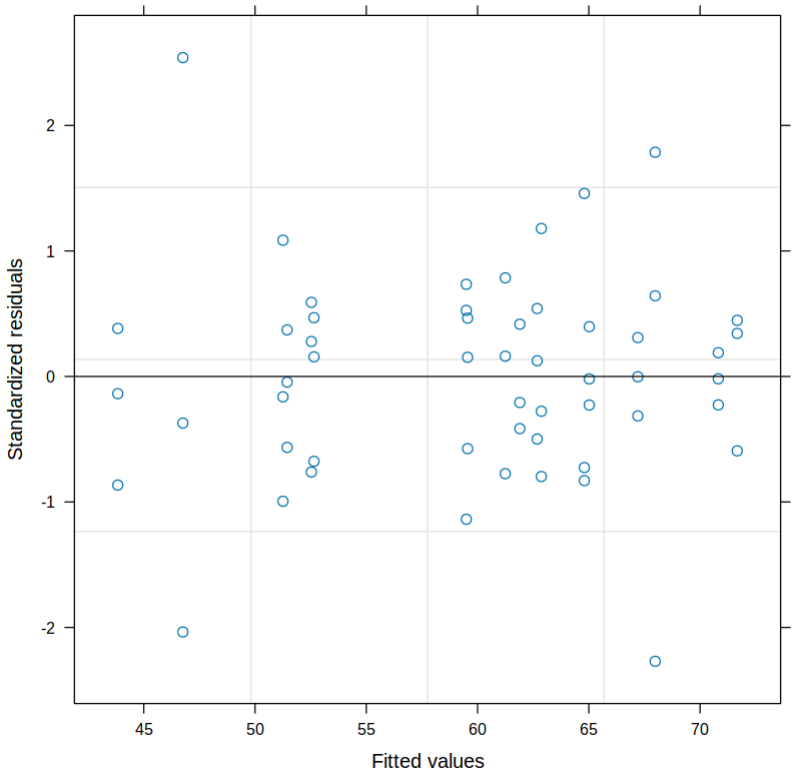


construimos un modelo con interaccion

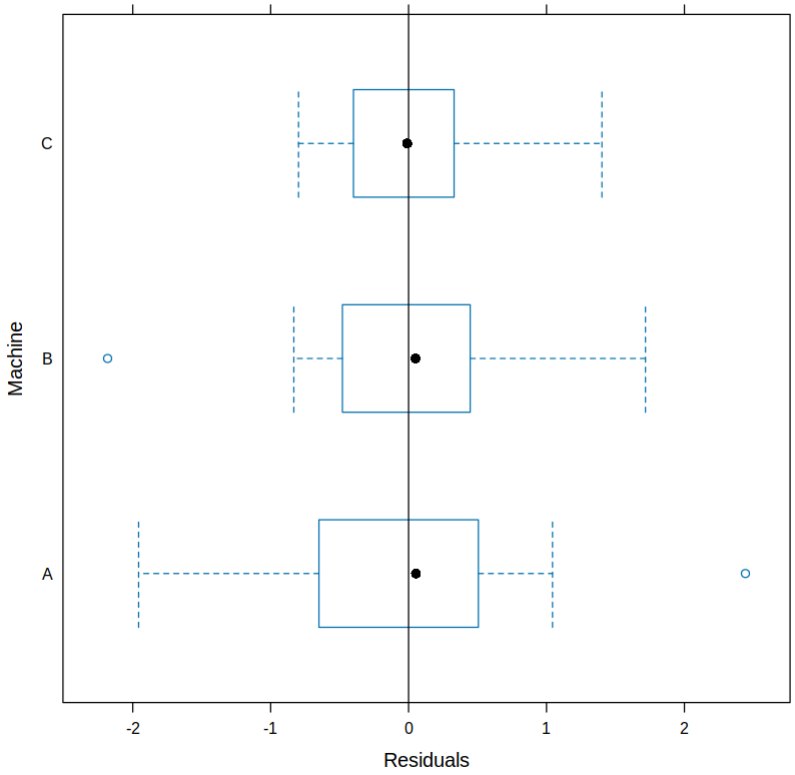
In [6]:
```
m1 <- lme( score ~ Machine, data = Machines, random = ~ 1 | Worker/Machine )
m1
```

```
Linear mixed-effects model fit by REML
  Data: Machines
  Log-restricted-likelihood: -107.8438
  Fixed: score ~ Machine
(Intercept)    MachineB    MachineC
  52.355556    7.966667   13.916667

Random effects:
 Formula: ~1 | Worker
        (Intercept)
StdDev:     4.78105

 Formula: ~1 | Machine %in% Worker
        (Intercept)  Residual
StdDev:    3.729532 0.9615771

Number of Observations: 54
Number of Groups:
             Worker Machine %in% Worker
                  6                  18
```

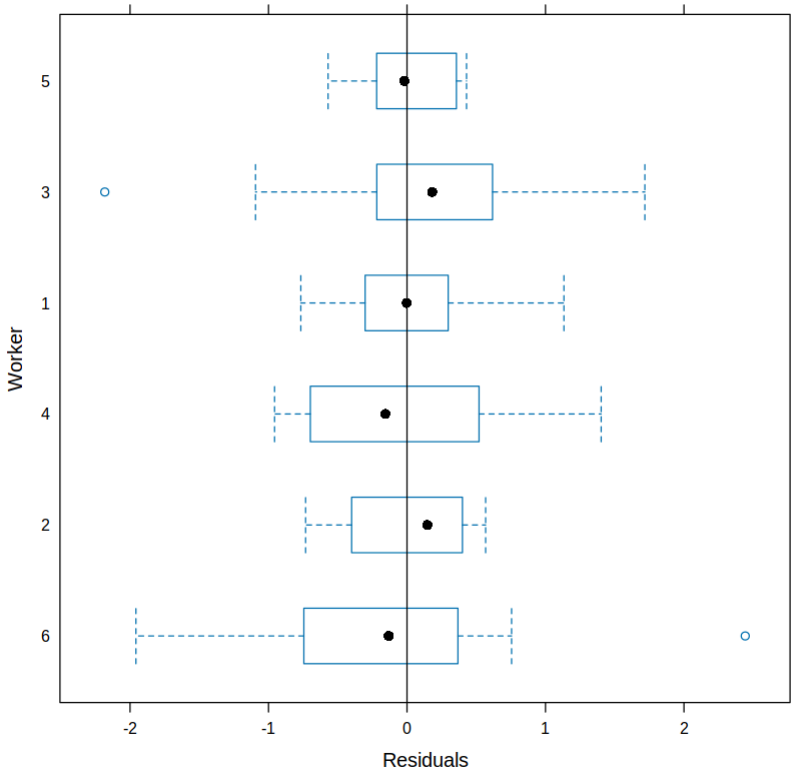
In [7]:
```
plot(m1)
```



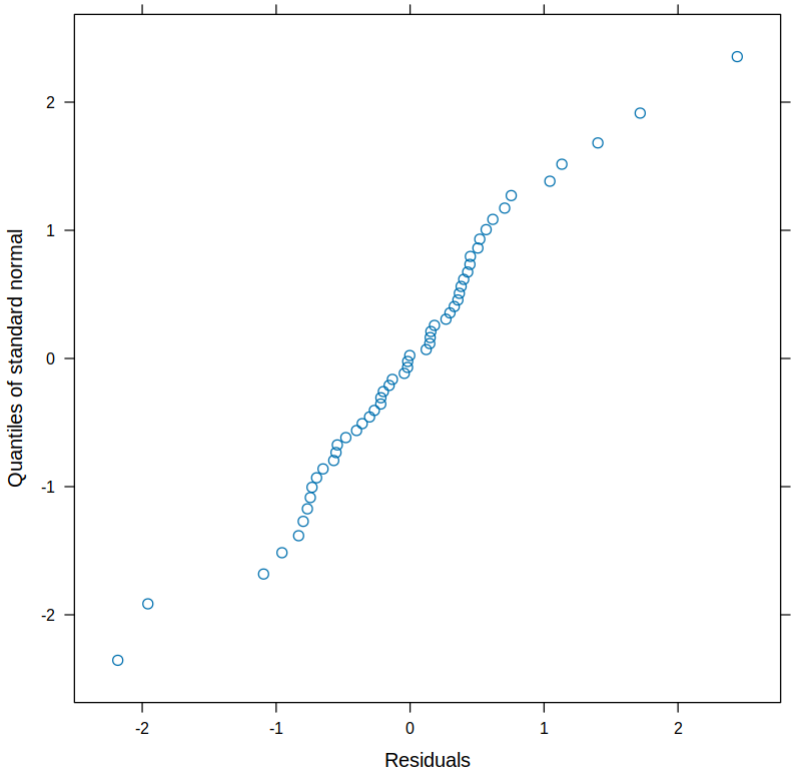
In [8]:
```
plot(m1,Machine~resid(.),abline=0)
```



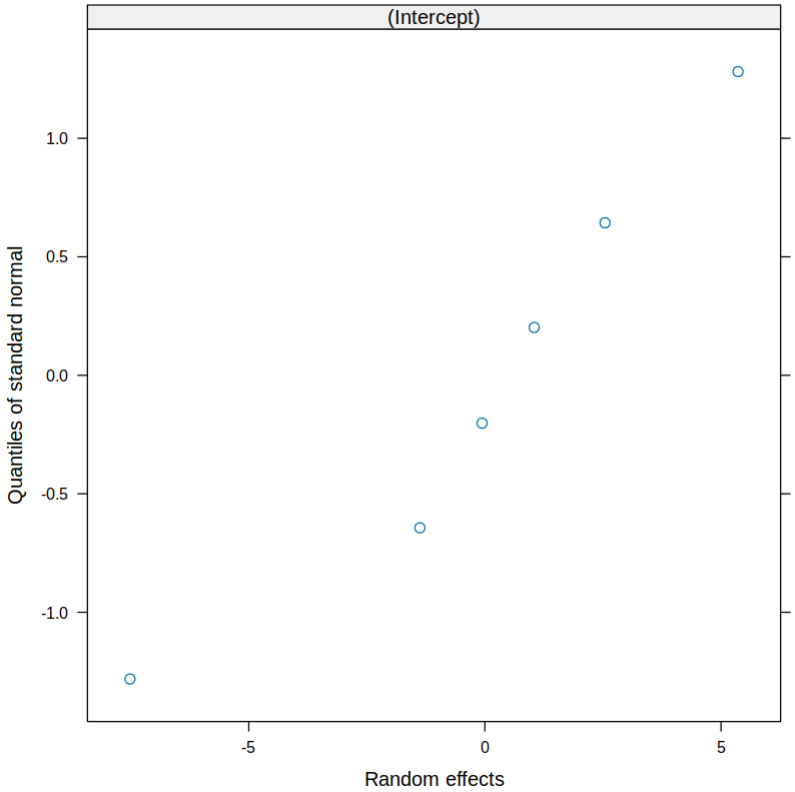
In [9]:
```
plot(m1,Worker~resid(.),abline=0)
```



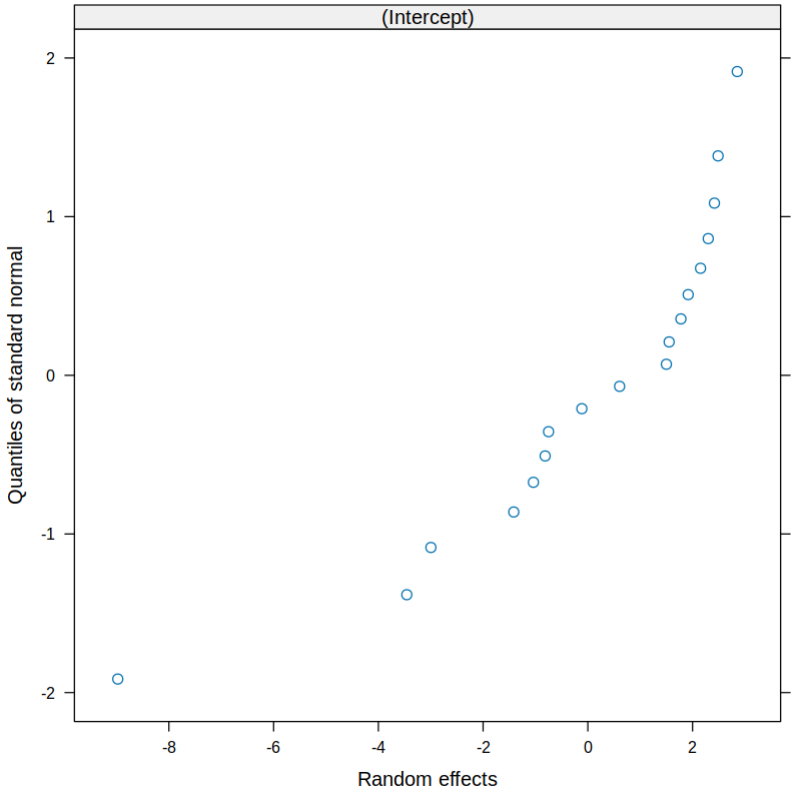
In [10]:
```
qqnorm(m1,~resid(.))
```



In [11]:
```
qqnorm(m1,~ranef(.,level=1))
```



In [12]:
```
qqnorm(m1,~ranef(.,level=2))
```



Se puede observar un dato atipico

(b) Asegúrese de probar si la interacción de la ec. 4 es apropiada.

In [13]:
```
anova(m0, m1)
```

A anova.lme: 2 × 9

	call	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-va
	<chr>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<fct>	<dbl>	<d
m0	lme.formula(fixed = score ~ Machine, data = Machines, random = ~1 | Worker)	1	5	296.8782	306.5373	-143.4391		NA	
m1	lme.formula(fixed = score ~ Machine, data = Machines, random = ~1 | Worker/Machine)	2	6	227.6876	239.2785	-107.8438	1 vs 2	71.19063	3.243

In [14]:
```
anova(m1)
```

A anova.lme: 2 × 4

	numDF	denDF	F-value	p-value
	<int>	<dbl>	<dbl>	<dbl>
(Intercept)	1	36	773.57093	0.0000000000
Machine	2	10	20.57615	0.0002855446

In [15]:
```
intervals(m1)
```

```
Approximate 95% confidence intervals

 Fixed effects:
               lower      est.    upper
(Intercept) 47.314062 52.355556 57.39705
MachineB     3.116071  7.966667 12.81726
MachineC     9.066071 13.916667 18.76726

 Random Effects:
  Level: Worker
                   lower    est.    upper
sd((Intercept)) 2.249875 4.78105 10.15987
  Level: Machine
                   lower     est.   upper
sd((Intercept)) 2.382833 3.729532 5.83734

 Within-group standard error:
   lower      est.     upper
0.7635351 0.9615771 1.2109864
```

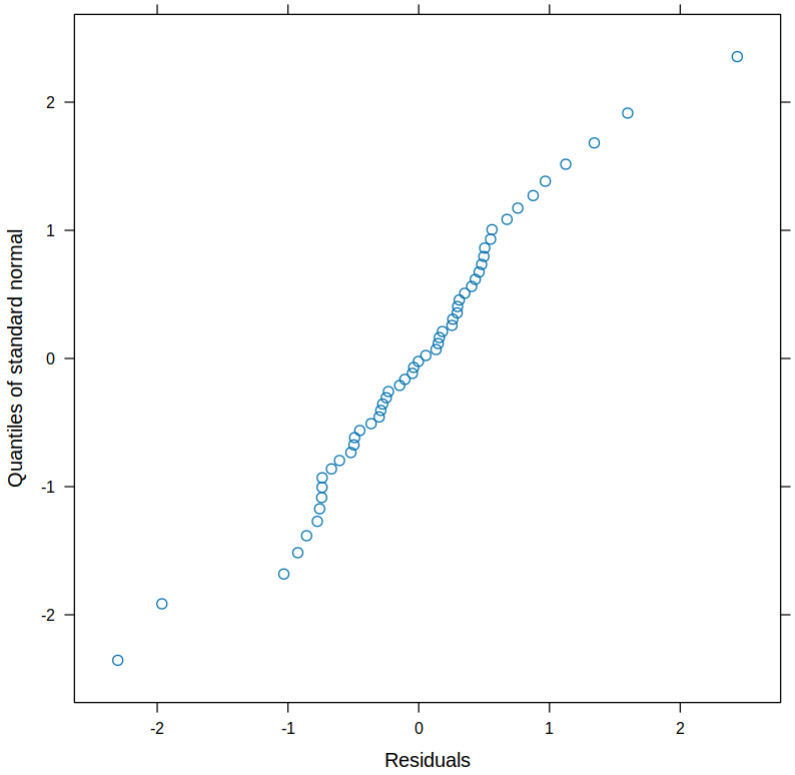
Todos los criterios dan preferencia al modelo con interaccion. Particularmente, el criterio con interaccion incrementa la verosimilitud.

(c) De manera similar, pruebe si sería apropiada una estructura de efectos aleatorios más compleja: específicamente una en la que la interacción máquina-trabajador esté correlacionada con la del trabajador.

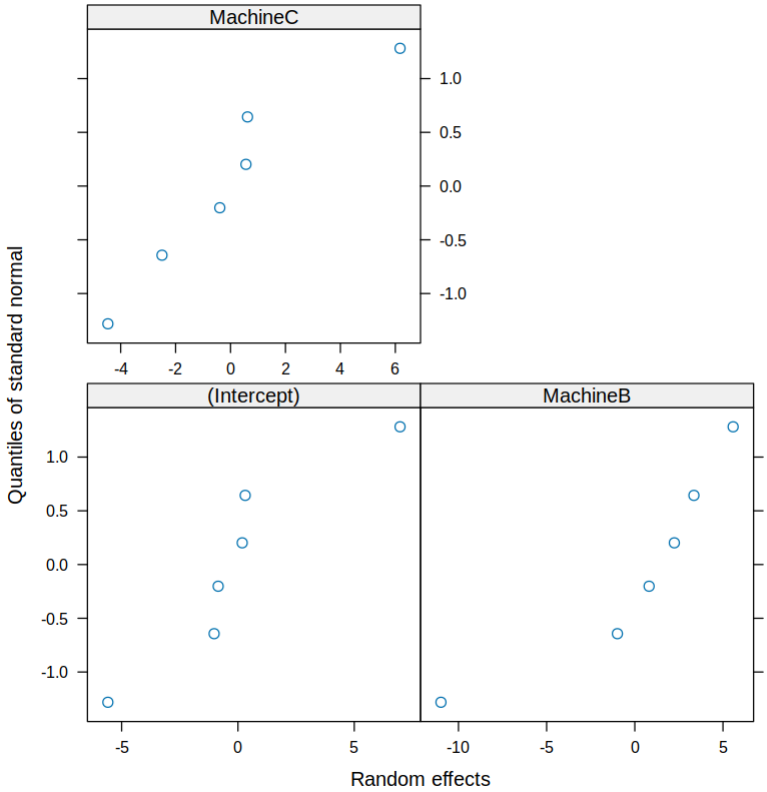
In [16]:
```
m2 <- lme( score ~ Machine, data = Machines, random = ~ Machine | Worker)
summary(m2)
```

```
Linear mixed-effects model fit by REML
  Data: Machines
       AIC      BIC    logLik
  228.3112 247.6295 -104.1556

Random effects:
 Formula: ~Machine | Worker
 Structure: General positive-definite, Log-Cholesky parametrization
            StdDev    Corr
(Intercept) 4.0792806 (Intr) MachnB
MachineB    5.8776433  0.484
MachineC    3.6898543 -0.365  0.297
Residual    0.9615766

Fixed effects:  score ~ Machine
               Value Std.Error DF   t-value p-value
(Intercept) 52.35556  1.680711 46 31.150834  0.0000
MachineB     7.96667  2.420851 46  3.290854  0.0019
MachineC    13.91667  1.540100 46  9.036211  0.0000
 Correlation:
         (Intr) MachnB
MachineB  0.463
MachineC -0.374  0.301

Standardized Within-Group Residuals:
        Min          Q1         Med          Q3         Max
-2.39354008 -0.51377574  0.02690829  0.47245471  2.53338699

Number of Observations: 54
Number of Groups: 6
```

In [17]:
```
qqnorm(m2,~resid(.))
```



In [18]:
```
qqnorm(m2,~ranef(.,level=1))
```



In [19]:
```
anova(m0,m1,m2)
```

A anova.lme: 3 × 9

	call	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-va
	<chr>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<fct>	<dbl>	<d
m0	lme.formula(fixed = score ~ Machine, data = Machines, random = ~1 | Worker)	1	5	296.8782	306.5373	-143.4391		NA	
m1	lme.formula(fixed = score ~ Machine, data = Machines, random = ~1 | Worker/Machine)	2	6	227.6876	239.2785	-107.8438	1 vs 2	71.19063	3.243
m2	lme.formula(fixed = score ~ Machine, data = Machines, random = ~Machine | Worker)	3	10	228.3112	247.6295	-104.1556	2 vs 3	7.37635	1.172

In [20]:
```
anova(m2)
```

A anova.lme: 2 × 4

	numDF	denDF	F-value	p-value
	<int>	<dbl>	<dbl>	<dbl>
(Intercept)	1	46	2351.80626	0.000000e+00
Machine	2	46	41.00377	5.984291e-11

(d) Si algún dato parece particularmente problemático en los gráficos de verificación, repita el análisis y vea si las conclusiones cambian.

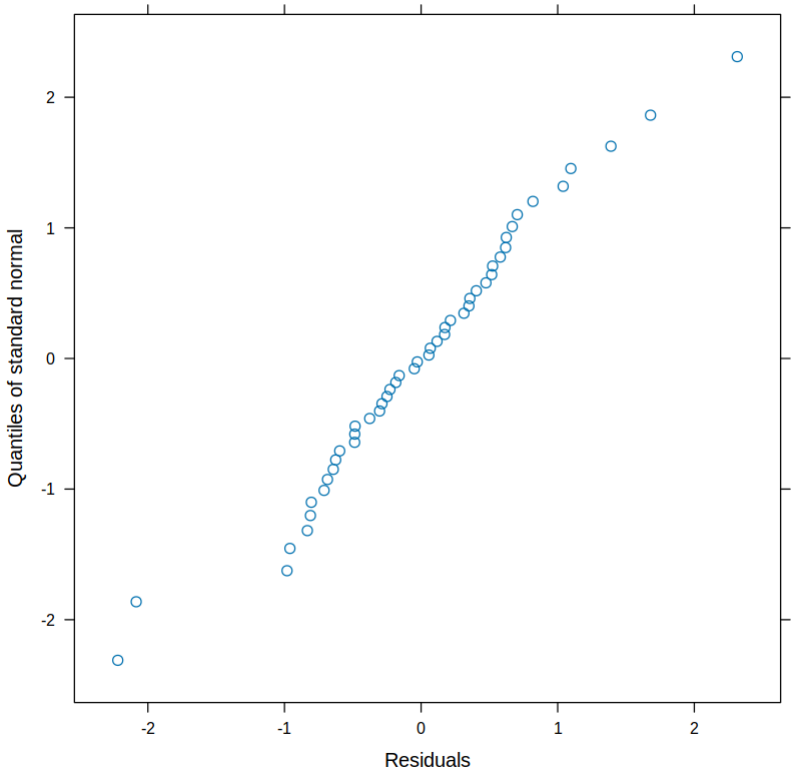
Se eliminan los datos relacionados con el trabajador 6

In [27]:
```
Machines <- Machines[-(34:36),]
```

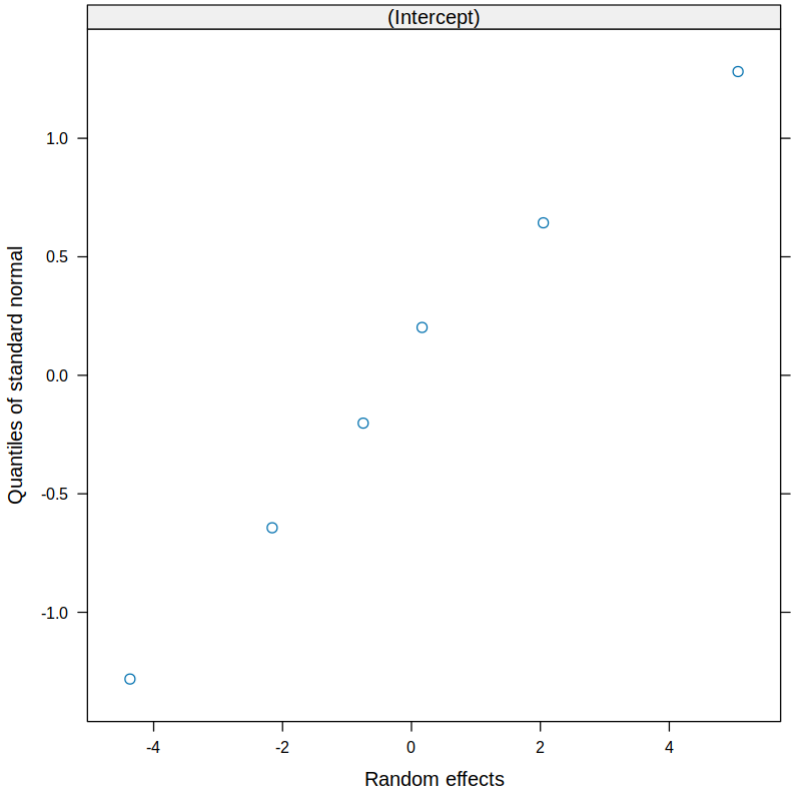
In [28]:
```
m3 <- lme( score ~ Machine, data = Machines, random = ~ 1 | Worker/Machine )
summary(m3)
```

```
Linear mixed-effects model fit by REML
  Data: Machines
       AIC      BIC    logLik
  193.0094 203.8493 -90.50468

Random effects:
 Formula: ~1 | Worker
        (Intercept)
StdDev:    3.530511

 Formula: ~1 | Machine %in% Worker
        (Intercept) Residual
StdDev:    2.084565  1.00592

Fixed effects:  score ~ Machine
               Value Std.Error DF   t-value p-value
(Intercept) 52.35556  1.690523 32 30.970030   0e+00
MachineB    10.43100  1.335769  8  7.808985   1e-04
MachineC    13.76405  1.335769  8 10.304214   0e+00
 Correlation:
         (Intr) MachnB
MachineB -0.346
MachineC -0.346  0.427

Standardized Within-Group Residuals:
        Min          Q1         Med          Q3         Max
-2.20732393 -0.51073470  0.01452005  0.51545659  2.30078469

Number of Observations: 48
Number of Groups:
             Worker Machine %in% Worker
                  6                  16
```

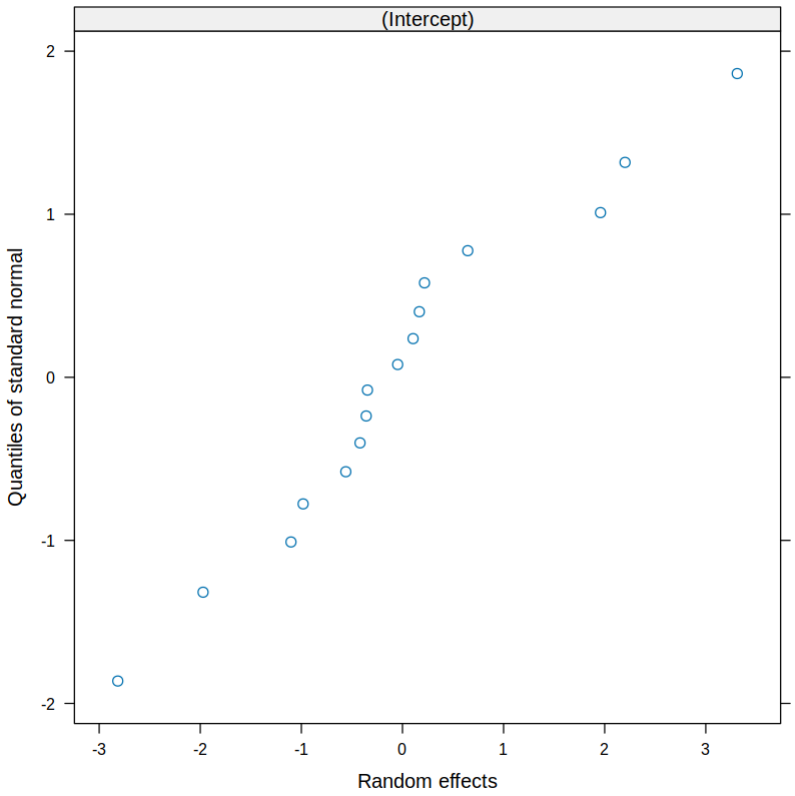
In [32]:
```
qqnorm(m3,~resid(.))
```



In [29]:
```
qqnorm(m3,~ranef(.,level=1))
```



In [30]:
```
qqnorm(m3,~ranef(.,level=2))
```


(Intercept)
Quantiles of standard normal
Random effects
2
1
0
-1
-2
-3
-2
-1
0
1
2
3