

P A R T E 2

El papel probabilístico normal y los diagramas de puntos en el análisis de la significación de los efectos

CAPÍTULO 2

El papel probabilístico normal y los diagramas de puntos en el análisis de la significación de los efectos

2.1 Antecedentes

La representación de los efectos en **papel probabilístico normal** (ppn) para analizar su significación estadística (método propuesto por Daniel en 1959) es en muchos ambientes no sólo el método más utilizado sino también “el método” por excelencia. Así es recogido en libros clásicos en el ámbito académico como el de Box, Hunter y Hunter (1978) o el de Montgomery (1997). Aunque la “Escuela de Taguchi” no los utiliza, ni tampoco libros de divulgación y libros orientados a la práctica, como el de Barrentine (1999) que utiliza el denominado “diagrama de Pareto de los efectos”, opción que también está implementada en el paquete estadístico Minitab.

El uso de esta técnica requiere tener claros conceptos estadísticos que si no se han entendido bien, o se olvidan, la acaban convirtiendo en una receta del tipo “los puntos que se separan de la recta son los significativos”, rodeada de cierto halo de misterio, y con la que se pueden cometer fácilmente “errores de bulto”.

Se conjetura que el uso de un simple diagrama de puntos de los efectos proporciona prácticamente la misma información que su representación en papel probabilístico normal, con la ventaja de que es mucho más fácil de entender y de recordar. La Figura 2.1a contiene una representación de los efectos en papel probabilístico normal y en diagrama de puntos del ejemplo

de moldeo por inyección que hacen Box, Hunter y Hunter (1993). En ambos diagramas resultan evidentes los efectos que se pueden considerar como significativos, no obstante alguien preocupado por la receta “los puntos que se separan de la recta son los significativos” puede confundirse con la ubicación que tiene que hacer de la recta y llegar a considerar significativos efectos que no lo son (Figura 2.1b), o depreciar efectos significativos (Figura 2.1c).

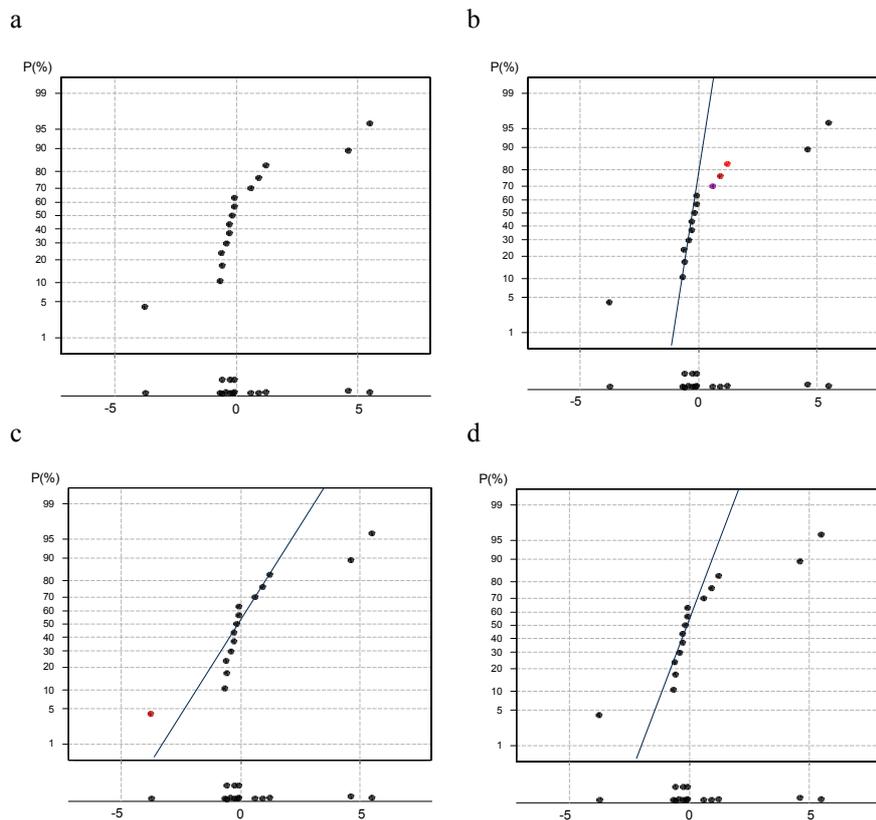


Figura 2.1: Representación de los efectos en papel probabilístico normal y en diagrama de puntos. Los casos b, c y d presentan distintas posibilidades de ubicar la recta “a mano”

En esta parte se presenta el estudio empírico que se realizó para analizar si esta conjetura es cierta; se describe la metodología empleada para la obtención de los datos, el análisis efectuado y las conclusiones obtenidas.

2.2 Plan de trabajo

Se realizó un estudio con alumnos de la Escuela Técnica Superior de Ingeniería Industrial de Barcelona. Se utilizó un cuestionario con 16 gráficos de efectos en ppn, 8 de ellos correspondientes a un diseño con 16 experimentos y otros 8 con 8 experimentos. El mismo cuestionario incluye 16 diagramas de puntos que representan los efectos obtenidos en otros tantos diseños experimentales, aparentemente distintos de los representados en ppn pero que en realidad corresponden a las mismas situaciones. Se pide a los estudiantes que en cada uno de los 32 gráficos presentados identifiquen los efectos que consideren significativos.

Las conclusiones del estudio se obtienen comparando las respuestas obtenidas, con sus “aciertos” y “fallos”.

2.2.1 Estudiantes con los que se realiza el estudio

El cuestionario se pasó en la última clase de la asignatura “Métodos Estadísticos de la Ingeniería I”. Esta asignatura es la primera que se cursa sobre estadística y está situada en el 5º semestre del plan de estudios. El programa consta de una presentación de los conceptos básicos (distribuciones de probabilidad, distribución de estadísticos muestrales, contraste de hipótesis, intervalos de confianza,...) comparación de 2 y más de 2 tratamientos y diseños factoriales.

El texto que se sigue es el de Prat et al. (1997) hasta el capítulo 8, que es el que trata sobre diseños factoriales fraccionales, aunque este capítulo no se estudia completo. El contenido de lo que se explica sobre diseños factoriales sigue el esquema: Importancia de la experimentación, ventajas de los diseños factoriales, planteamiento de este tipo de diseños, cálculo de los

efectos, análisis de su significación estadísticas (cuando hay réplicas a través de la estimación de la varianza de los efectos, y cuando no las hay utilizando ppn), interpretación de los resultados obtenidos, estrategia secuencial y diseños fraccionales (generador del diseño, relación de definición, deducción de la estructura de alias y concepto de resolución del diseño). Este contenido corresponde aproximadamente con los capítulos 10 y 12 del texto de Box, Hunter y Hunter (1993) o el capítulo 12 del de Montgomery (2001).

El diseño de experimentos es el último tema que se trata en el curso; por tanto, el cuestionario se pasó cuando el tema estaba todavía fresco, aunque seguramente no se había estudiado para el examen final. Hay que destacar que no se explica el uso del diagrama de puntos para discriminar los efectos significativos.

2.2.2 Diseño del cuestionario

El cuestionario, tal como se presentó a los estudiantes, se encuentra en el Anexo 4. Como se ha comentado presenta 32 gráficos y en 16 de ellos se utiliza el ppn y en los otros 16 el diagrama de puntos, aunque en realidad sólo existen 16 situaciones distintas, cada una de ellas representada con los 2 tipos de gráfico. Esto facilita la comparación, aunque es importante que los estudiantes no se den cuenta de este hecho mientras están identificando los efectos significativos, ya que su deseo de ser coherentes invalidaría el estudio. La disposición de los gráficos en el cuestionario hace que esta característica sea difícil de apreciar, especialmente cuando se dispone de un tiempo reducido para contestar y uno se va concentrando en el análisis caso a caso.

2.2.3 Selección de los casos

Una posibilidad era inventar los valores de los efectos en los 16 casos, intentando presentar situaciones que fueran representativas. Sin embargo este sistema, aunque fácil y rápido, tiene el inconveniente de que se puede

tender a plantear situaciones demasiado claras, favoreciendo así la igualdad de los 2 métodos de análisis. Por otro lado, en situaciones poco claras es discutible lo que debe considerarse significativo y lo que no, lo que complicaría la comparación.

Lo que se hizo fue plantear un modelo y generar de forma aleatoria los coeficientes y el término del error para cada uno de los casos. A partir de la ecuación obtenida se calcularon las respuestas, y a partir de las respuestas los efectos, lo que permite tener el efecto real y el efecto estimado. Para diseños con 16 experimentos se tomó el modelo:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \varepsilon$$

Para diseños con 8 experimentos, el modelo es el mismo con $x_4 = 0$. Los coeficientes y el valor del error se generaron de acuerdo con las distribuciones mostradas en la Tabla 2.1.

Tabla 2.1 Distribución de los parámetros del modelo

| Coefficientes | Distribuciones |
|---------------|---|
| β_0 | U(0,100) |
| β_1 | Ber(0.7)*U(0,20) |
| β_2 | Ber(0.7)*U(10,30) |
| β_3 | Ber(0.7)*U(20,40) |
| β_4 | Ber(0.7)*U(15,35) |
| β_{12} | $\beta_1 * \beta_2 * \text{Ber}(0.5) * U(0,20) / 200$ |
| β_{13} | $\beta_1 * \beta_3 * \text{Ber}(0.5) * U(0,20) / 300$ |
| β_{14} | $\beta_1 * \beta_4 * \text{Ber}(0.5) * U(0,20) / 300$ |
| β_{23} | $\beta_2 * \beta_3 * \text{Ber}(0.5) * U(0,20) / 600$ |
| β_{24} | $\beta_2 * \beta_4 * \text{Ber}(0.5) * U(0,20) / 500$ |
| β_{34} | $\beta_3 * \beta_4 * \text{Ber}(0.5) * U(0,20) / 600$ |
| σ | U(4,8) |
| ε | N(0, σ) |

Todo este proceso se realizó utilizando el paquete estadístico Minitab y dos macros, una para los diseños 2^3 y otra para los 2^4 . La macro para este último caso se encuentra en el Anexo 5.

Aunque la primera idea era generar 16 situaciones e incluirlas en el cuestionario tal como salieran, hecho de esta forma podían aparecer gráficos de interpretación evidente y también aparecer varios prácticamente repetidos. Para evitar que se diera esta circunstancia se generaron 20 situaciones para cada tipo de diseño, y de cada una de ellas elegimos las 8 que nos parecieron más convenientes, tratando de abarcar diferentes niveles de dificultad. Estas situaciones se presentan en el Anexo 6.

2.2.4 Disposición de los casos en el cuestionario

El cuestionario cuenta con 4 hojas, y cada una de ellas incluye 8 situaciones, 4 en ppn y otras 4 en diagrama de puntos, de forma que en una misma hoja no hay situaciones repetidas. Ver Figura 2.2.

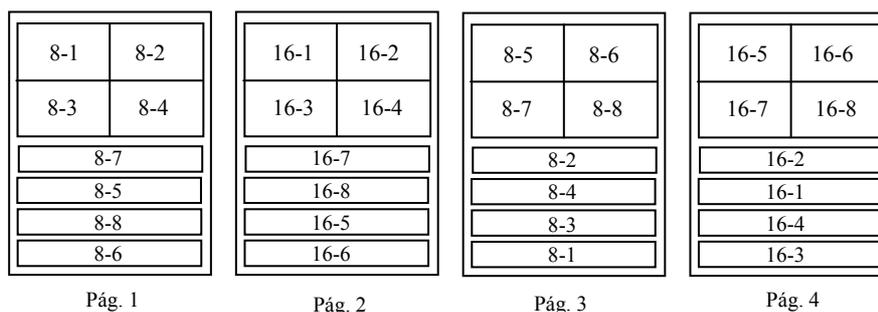


Figura 2.2 Disposición de las situaciones que se presentan en las 4 páginas del cuestionario. El primer valor se refiere al número de experimentos y el segundo al número del caso

2.2.5 Pase del cuestionario

Se realizó un primer pase del cuestionario en la última clase de la asignatura impartida en el primer semestre del curso 2001-2002. El cuestionario se entregó 15 minutos antes de acabar la clase, por lo que no se indicó un

tiempo máximo para contestar, aunque ya estaba claro que disponían de un máximo de 15 minutos. No se dio ninguna instrucción ni se hizo ningún comentario adicional a los que figuran en la primera página del cuestionario.

Este primer pase se realizó en 2 grupos (2 clases distintas) obteniendo un total de 47 cuestionarios rellenados correctamente. La intención fue utilizar este primer pase como encuesta piloto para depurar los posibles fallos o malas interpretaciones no previstas. Sin embargo, después de analizar los resultados se llegó a la conclusión de que no era necesaria ninguna corrección. El único problema observado era que algunos estudiantes dejaban algún caso en blanco y no estaba claro si había sido por falta de tiempo, por no saber que efectos marcar como significativos, o por considerar que ninguno lo era. Como los que contenían casos en blanco sólo eran 3 decidimos dejarlo como estaba y repetir la encuesta exactamente igual la próxima vez que se dictara la asignatura. Los cuestionarios que tenían casos en blanco se eliminaron del estudio.

El segundo pase se hizo también en 2 grupos, exactamente igual que el primero, y se obtuvieron 69 cuestionarios correctos (se eliminaron 2 por tener casos en blanco). Posteriormente se hizo un tercer pase con cuatro grupos y se aumentaron 100 cuestionarios, de los cuales se eliminaron 5.

En total disponemos de 211 cuestionarios considerados correctos (se eliminaron 10 por contener casos en blanco), y aunque fueron recogidos en 2 cursos académicos y en un total de 8 grupos distintos, se han analizado como un conjunto homogéneo ya que tanto el programa como los profesores y las otras condiciones relevantes fueron las mismas en todos los casos.

2.2.6 Resultados obtenidos

Se han considerado significativos los efectos cuyo valor real es distinto de cero, con independencia del aspecto que presente el gráfico. Los resultados totales están detallados, caso a caso, en el Anexo 6.

Tabla 2.2 Totales de puntos identificados y marcados erróneamente por casos planteados

| Casos planteados | Efectos significativos del caso | Efectos significativos en todos los cuestionarios (211) | Marcados | | Faltan | | Sobran | | Errores (Faltan + Sobran) | |
|------------------|---------------------------------|---|-------------|--------------------|-------------|--------------------|-------------|--------------------|---------------------------|--------------------|
| | | | Gráfico ppn | Diagrama de puntos | Gráfico ppn | Diagrama de puntos | Gráfico ppn | Diagrama de puntos | Gráfico ppn | Diagrama de puntos |
| 2 ³ | 1 | 4 | 456 | 431 | 388 | 413 | 0 | 0 | 388 | 413 |
| | 2 | 3 | 647 | 586 | 88 | 105 | 102 | 58 | 190 | 163 |
| | 3 | 5 | 443 | 609 | 615 | 448 | 3 | 2 | 618 | 450 |
| | 4 | 4 | 650 | 675 | 208 | 170 | 14 | 1 | 222 | 171 |
| | 5 | 2 | 424 | 428 | 6 | 4 | 8 | 10 | 14 | 14 |
| | 6 | 1 | 368 | 470 | 10 | 14 | 167 | 273 | 177 | 283 |
| | 7 | 2 | 337 | 380 | 149 | 119 | 64 | 77 | 213 | 196 |
| | 8 | 3 | 462 | 533 | 227 | 166 | 56 | 66 | 283 | 232 |
| 2 ⁴ | 1 | 4 | 803 | 766 | 80 | 99 | 39 | 21 | 119 | 120 |
| | 2 | 5 | 1007 | 1002 | 52 | 60 | 4 | 7 | 56 | 67 |
| | 3 | 7 | 934 | 922 | 555 | 566 | 12 | 11 | 567 | 567 |
| | 4 | 4 | 876 | 713 | 111 | 228 | 143 | 97 | 254 | 325 |
| | 5 | 3 | 436 | 610 | 365 | 292 | 168 | 269 | 533 | 561 |
| | 6 | 8 | 1391 | 1146 | 299 | 542 | 2 | 0 | 301 | 542 |
| | 7 | 4 | 748 | 755 | 156 | 150 | 60 | 61 | 216 | 211 |
| | 8 | 7 | 826 | 810 | 656 | 679 | 4 | 14 | 660 | 693 |

La Tabla 2.2 resume el número de efectos que se identificaron correcta e incorrectamente. La columna “Efectos significativos en todos los cuestionarios (211)” se ha obtenido multiplicando la columna anterior (número de efectos significativos en cada caso) por 211, que es el número de cuestionarios total. Bajo el encabezado “Marcados” se tiene el número de efectos que se marcaron como significativos, independientemente de que lo sean o no; en “Faltan” se indica el número de efectos significativos que no se marcaron y en “Sobran” se tiene el número de los que se marcaron como significativos sin serlo.

Para cada método, el número de efectos significativos en los 211 cuestionarios es igual a la suma de los efectos marcados, más los que faltan, menos los que sobran.

2.3 Análisis de los resultados

La Figura 2.3 muestra el número de efectos significativos que no han sido identificados (faltan), los efectos marcados significativos sin serlo (sobran) y el total de errores, según el método utilizado, para cada una de las situaciones planteadas y de acuerdo al número de pruebas en el experimento. Con ambos métodos se cometen más equivocaciones al no identificar efectos significativos que las que se dan al marcar efectos que no lo son.

Para el total de errores y 8 experimentos existe una ligera tendencia a producirse más equivocaciones con el papel probabilístico normal que con el diagrama de puntos. Esta tendencia se invierte cuando se tienen 16 experimentos.

Aparecen 3 casos¹ (los 8-3, 8-6 y 16-6) en que las diferencias son importantes comparadas con el resto. Los casos 16-4 y 16-5 sobresalen en las situaciones “faltan” y en “sobran” pero prácticamente las diferencias se

¹ La notación 8-3 significa el caso 3 con 8 experimentos

anulan al considerar el total de errores, esto es así porque en una situación resulta mejor un método y pasa lo contrario en la otra situación.

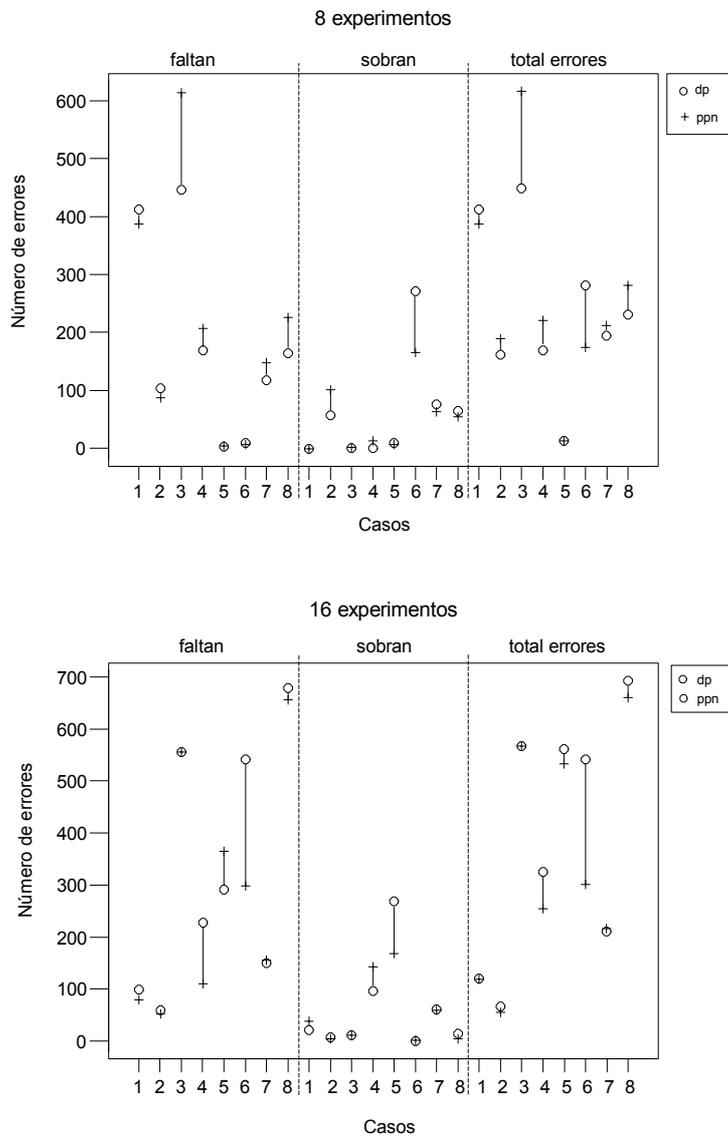


Figura 2.3 Número de errores debido a efectos significativos sin identificar (faltan), efectos marcados significativos sin serlo (sobran) y total de errores, según el método utilizado, para cada una de los casos y de acuerdo al número de experimentos.

En la Figura 2.4 se tiene los diagramas de puntos de las diferencias $ppn - dp$ para cada tipo de error y para los errores totales, en ellos se pone de manifiesto que existen dos tipos de valores: los que están agrupados en torno a cero, y los que representan claras discrepancias respecto a estas distribuciones (los casos señalados anteriormente).

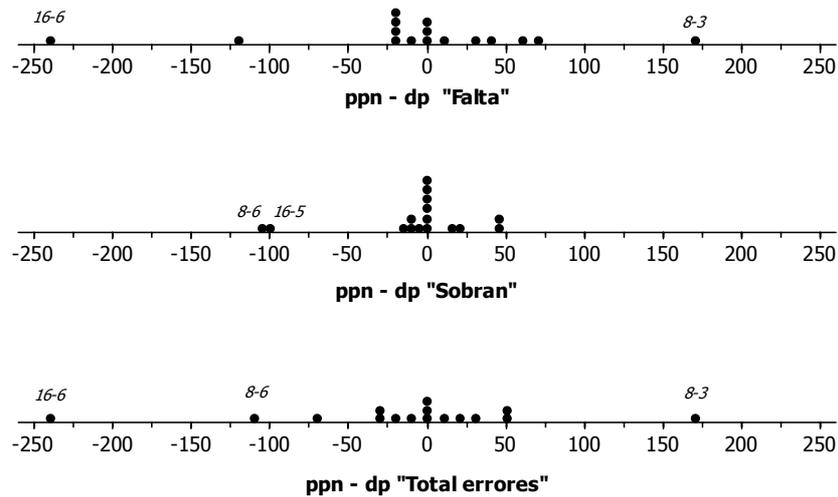


Figura 2.4 Diagrama de puntos de las diferencias entre los dos métodos, para cada situación de error considerada

Estas discrepancias parecen estar asociadas con el número de efectos significativos que tiene el caso, como se puede observar en la Figura 2.5 para el total de errores.

Aparentemente no funciona del todo bien el ppn en los experimentos con 8 pruebas cuando hay muchos efectos reales, como sucede en el caso 8-3 que tiene 5 efectos reales, quedando pocos puntos para definir la recta. La figura 2.6 reproduce los gráficos en ppn y dp de este caso, se nota que en estas circunstancias el uso del dp facilita la identificación de los efectos significativos.

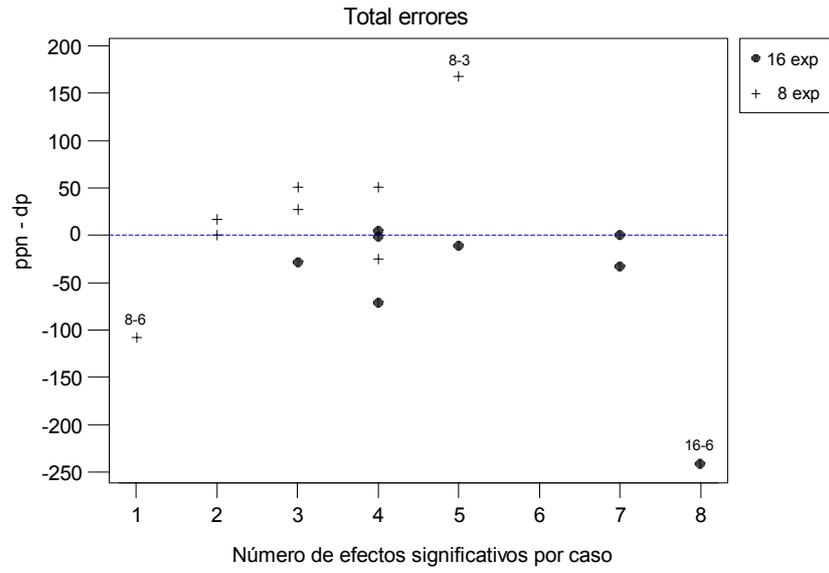


Figura 2.5 Relación entre las diferencias y el número de efectos significativos por caso

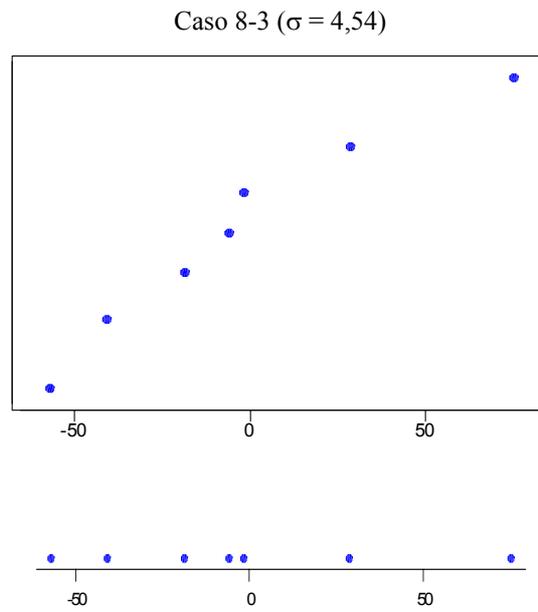


Figura 2.6 Representación de los efectos en ppn y en dp del caso 8-3 con 5 efectos significativos

Cuando se tienen más puntos para definir la recta el ppn funciona algo mejor que el diagrama de puntos, quedando evidente este hecho en la siguiente Figura de la discrepancia 16-6.

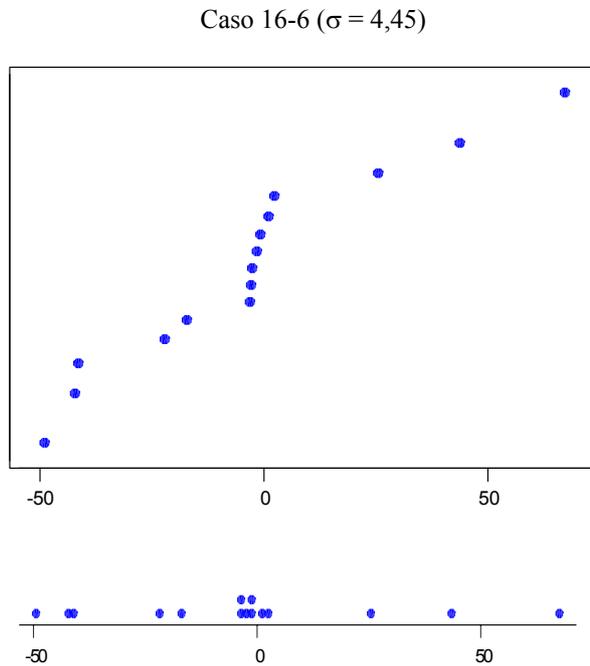


Figura 2.7 Representación de los efectos en ppn y en dp del caso 16-6 con 8 efectos significativos

2.4 Conclusiones

Del estudio realizado no puede concluirse que el método del ppn sea mejor que el uso de un simple diagrama de puntos de los efectos. Sin embargo, existen algunos casos en que el uso del diagrama de puntos es aparentemente ventajoso cuando se trata de analizar los efectos de diseños con 8 experimentos, y otros en que sí parece que el uso del ppn puede evitar errores en algunos casos con 16 experimentos. En el estudio realizado esto ocurre de forma clara en dos de los casos planteados para el total de errores.

Una interpretación que parece razonable es que mientras con 7 efectos es muy difícil apreciar la recta de los no significativos y por tanto el método de ppn ayuda poco, con 15 efectos ya se pueden tener fácilmente del orden de 10 no significativos, lo que sí puede ayudar a marcar una recta que suponga, en algunos casos, una ventaja respecto al diagrama de puntos.

Pero a la vista de que las diferencias no son muy importantes, y especialmente cuando uno de los objetivos sea animar al uso de los diseños factoriales yendo al meollo del asunto y despojándolos de sofisticaciones innecesarias, parece razonable considerar la oportunidad de utilizar el diagrama de puntos, sin duda más intuitivo y fácil de recordar. Baste mencionar que no se enseñó el uso del diagrama de puntos para discriminar los efectos significativos, los estudiantes contaban únicamente con la idea de que los efectos inertes toman valores diferentes de cero debido a la variabilidad inherente en todo proceso de experimentación y medición.

Una solución salomónica y fácil de aplicar es incluir un diagrama de puntos siempre que se utilice ppn, simplemente proyectando los puntos sobre el eje horizontal. Por supuesto, esta idea no es nuestra ya que, por ejemplo, así se presentan todas las representaciones en ppn en el texto de Box, Hunter y Hunter (1978), aunque no en otros textos también muy conocidos, ni tampoco en paquetes de software estadístico de amplia difusión como Minitab.

Aportaciones

Se efectuó un estudio empírico con un conjunto de alumnos para constatar que el uso de los diagramas de puntos puede ser una opción simple, frente al uso del papel probabilístico normal, para discriminar los efectos significativos en el análisis de experimentos con 8 o 16 pruebas.