UNA JUSTIFICACIÓN PROBABILÍSTICA DE LA DEMOCRACIA DIRECTA. PARTE II

Lex Kravetsky

ciber-comunismo

Lex Kravetsky es un programador y matemático que divulga abundante contenido sobre diferentes temas en su canal de YouTube (https://www.youtube.com/@KravetskiLex) y en diferentes blogs (https://lex-kravetski.livejournal.com/). La versión original del artículo en ruso puede consultarse en:https://22century.ru/popular-science-publications/direct-democracy-2 30 de agosto de 2019

El artículo es la segunda parte de una serie de tres.

¿Hasta qué punto es eficaz la democracia directa? ¿Es cierto que votando sólo los especialistas se resuelven mejor los problemas, que mediante el sufragio universal? Seguimos explorando estas cuestiones con ayuda de las matemáticas.

Eficacia de un grupo pequeño en comparación con un grupo grande

Hay otra conclusión filosófica "que se me pasa por la cabeza" sobre la democracia: "la democracia es buena, pero, por supuesto, sólo los especialistas en estos temas deberían votar sobre determinadas cuestiones".

Sin embargo, a pesar de su aparente intuitividad, también resulta ser errónea en el caso general: al fin y al cabo, cuanto menor sea el número de personas del colectivo, mayor será la probabilidad media de acertar de cada uno de sus miembros para conseguir la misma eficacia de acierto por parte de todo el colectivo.

Al seleccionar a un pequeño grupo de especialistas del colectivo que tomarán la decisión por votación, reducimos radicalmente el tamaño del colectivo y, por tanto, aumentamos radicalmente los requisitos de la probabilidad media de acertar entre los especialistas.

Sí, es posible encontrar condiciones en las que la identificación de un pequeño grupo de "expertos" con derecho a voto aumente la probabilidad de acertar la decisión correcta. Sin embargo, en la práctica esto no es posible en todos los casos: los requisitos para que la probabilidad de acertar entre este pequeño grupo sea comparable a la de un grupo grande pueden no ser alcanzables en absoluto en el mundo real.

Supongamos que decidimos sustituir un equipo de 100.000 personas por un equipo de 30 especialistas. En un equipo grande de no especialistas, la probabilidad de adivinar es de 0,51. ¿Cuál tendría que ser la probabilidad de acertar de los especialistas para ser tan buena como la del equipo?

Debe ser 0,93. Cada uno de los no especialistas puede acertar ligeramente mejor que la moneda, pero cada uno de los especialistas tendrá que acertar el 93% de las veces.

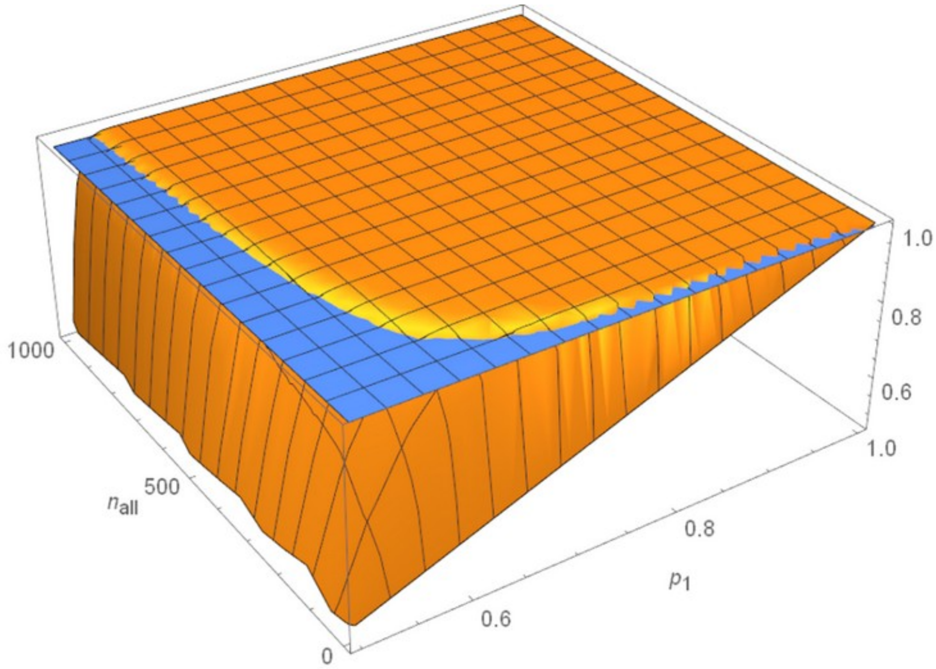
Pero, ¿qué ocurre si un equipo de no especialistas adivina con una probabilidad de 0,52? En este caso, cada especialista tiene que acertar con una probabilidad de 0,9989.

Si la probabilidad de acertar de los no especialistas es de 0,55, la probabilidad exigida a los especialistas debería diferir tan poco de 1, que en la práctica significa que "deben acertar siempre".

Por otra parte, si es posible elegir a 1.000 personas especialmente talentosas de entre 100.000 que adivinan con una probabilidad media de 0,51, puede estar justificado, porque los "especialistas" necesitan la probabilidad media de 0,60 para superar a este grupo. Esto es totalmente factible.

Sin embargo, esto se complica por el hecho de que, de nuevo, la probabilidad requerida de adivinar, para un equipo grande demasiado rápido se vuelve indistinguible de la unidad. Por ejemplo, si la probabilidad de acertar en el equipo anterior no es de 0,51, sino de 0,60, entonces un equipo de 1.000 "expertos" ya se vería obligado a tener una probabilidad media de acertar de 0,996 para cada uno de ellos.

En general, la dependencia de la probabilidad de tomar la decisión correcta para el colectivo depende de la probabilidad de adivinar de cada persona y del número de personas del colectivo tiene el siguiente aspecto (para mayor claridad, la zona por encima de la cual la probabilidad de tomar la decisión correcta colectivamente es superior a 0,99 está sombreada en azul).



Cuándo se sabe que la democracia es ineficaz

Del gráfico tridimensional anterior se desprende claramente que existe sin duda una zona en la que la probabilidad de tomar una decisión correcta es inferior a cualquier número dado en el intervalo (½, 1). Es decir, siempre podemos elegir un equipo con unos parámetros tales que sea peor que el dictador que nos han dado. Sin embargo, para ello, en la mayoría de los casos tendríamos que hacer que este equipo fuera muy pequeño y con una probabilidad media de acertar la solución superior a ½ muy baja.

Sin embargo, surge la pregunta: ¿existen parámetros que permitan tomar un dictador al azar del propio colectivo?

Esto equivale a preguntarse si existe una zona en la que

$$p_{dem}(n_{all}, p_1) < p_1$$

y al mismo tiempo

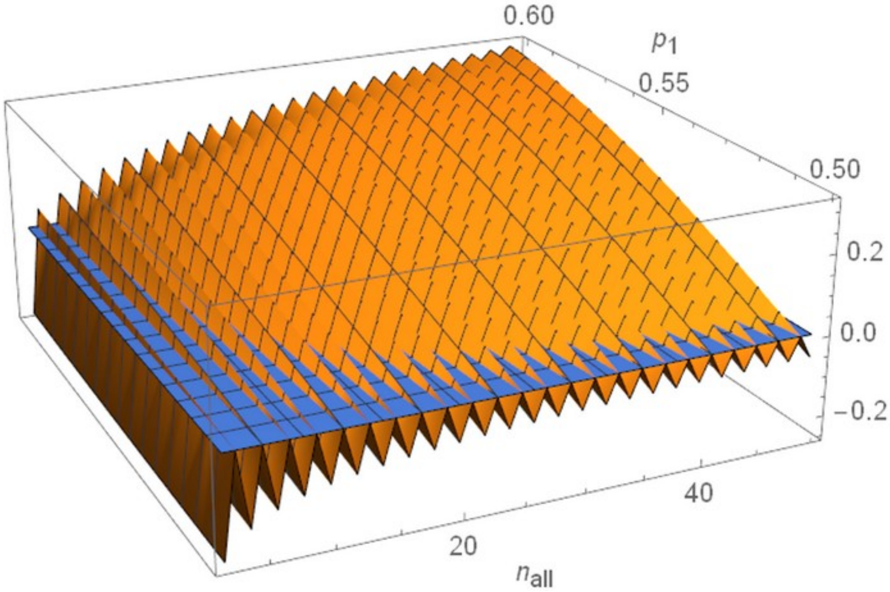
$$0{,}5 < p_1 < 1$$

$$n_{all} \geq 2$$

Para responder a esta pregunta, representa gráficamente la función

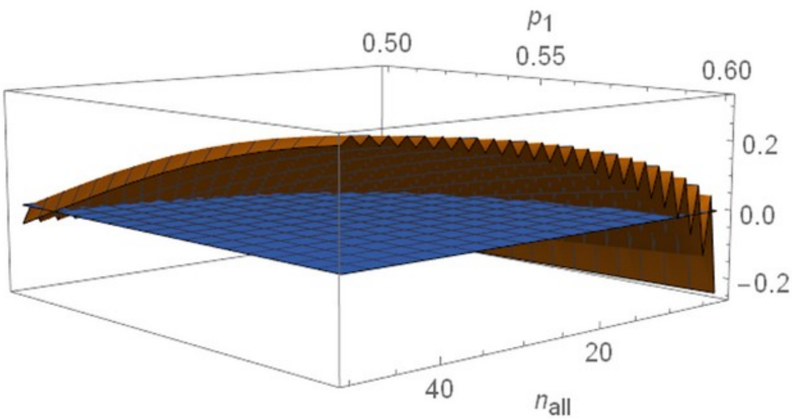
$$p_{dem}(n_{all}, p_1) - p_1$$

Cuando esta función es inferior a cero, un dictador elegido entre el colectivo es más ventajoso que tomar una decisión por votación (el plano azul corresponde aquí a cero).

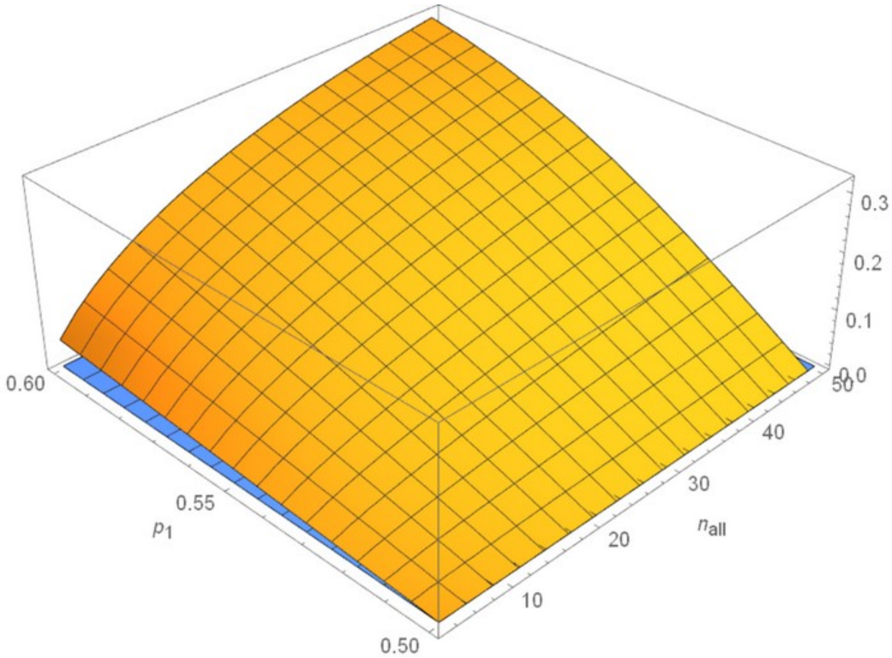


En este gráfico la "peineta" surge por las peculiaridades de las reglas del juego: a igualdad de votos se elige la opción equivocada, lo que lleva a que con un número par de participantes la probabilidad de acertar sea algo menor que con los impares vecinos.

Vemos que la diferencia que nos interesa es inferior a cero sólo cuando el número de participantes es par. E incluso esto ocurre sólo en la región del número pequeño de participantes y muy cerca de ½ probabilidad de acertar cada uno de ellos.



Si excluimos los números pares problemáticos de personal, entonces la función es positiva en todas partes



En otras palabras, sólo la confusión sobre qué hacer cuando el número de votos a favor y en contra resulta ser igual en un pequeño colectivo con un número par de participantes, en el que, por término medio, todos aciertan con una probabilidad cercana al resultado de lanzar una moneda al aire, podría permitir que un dictador elegido al azar entre el colectivo fuera más eficaz que una votación.

También en este caso es necesaria una aclaración.

Por supuesto, podemos tener la suerte de elegir al azar a alguien especialmente talentoso, que adivine mejor que nadie. Pero también podemos tener mala suerte y elegir a una persona que acierte peor que la mayoría de los demás.

Por lo tanto, la afirmación anterior debe entenderse de la siguiente manera: si elegimos muchas veces al azar dictadores de diferentes colectivos, entonces de media tomarán decisiones notablemente peores que este colectivo votando. Sólo cuando haya un número par pequeño de personas en el equipo, la probabilidad media de que cada una de ellas acierte sea cercana a ½, y las reglas digan que perderán la opción correcta por empate, los dictadores elegidos al azar darán, de media, un resultado mejor que ese equipo.

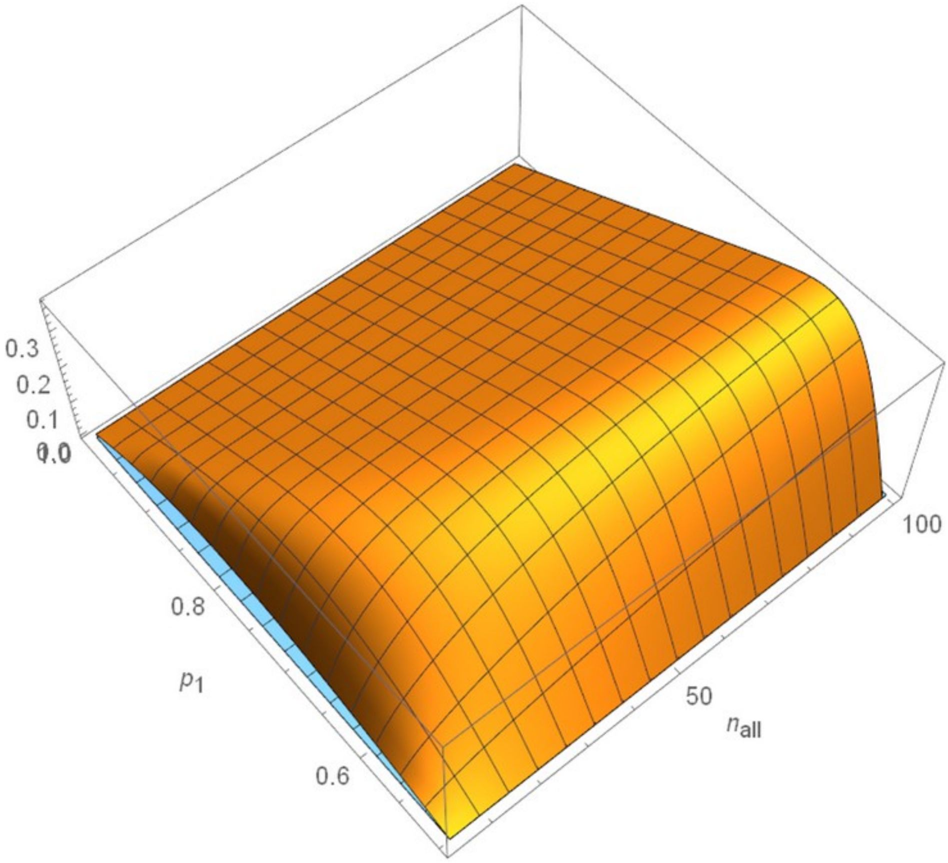
En este caso, la solución más sencilla −por ejemplo, excluir a un participante al azar cada vez para evitar un empate, o elegir la opción lanzando una moneda al aire en caso de empate− ya hará que un dictador nunca resulte ventajoso incluso en este caso.

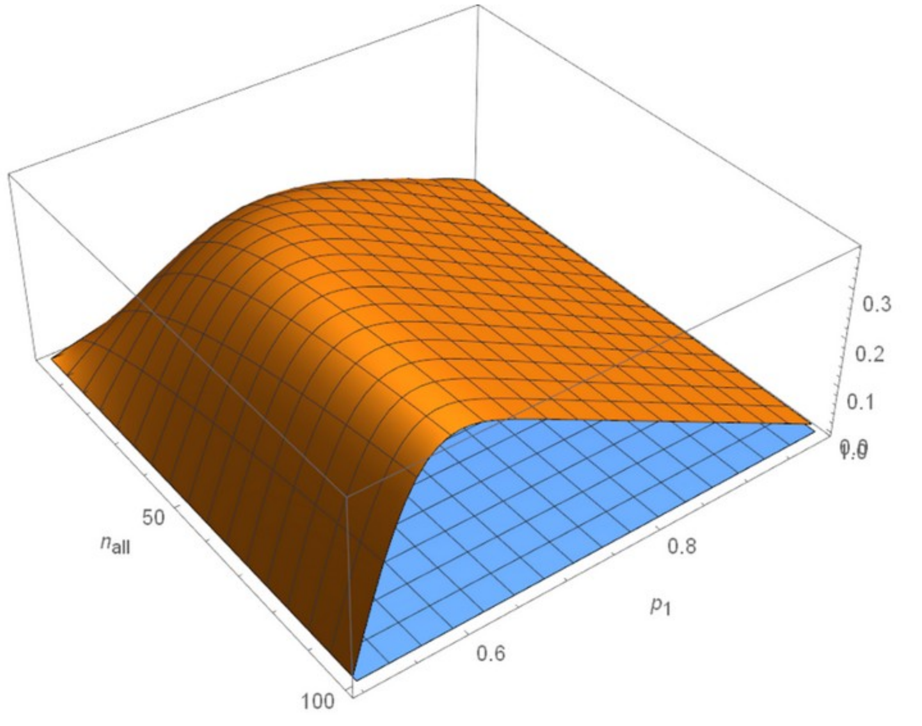
El gráfico de la ventaja de la democracia

La gráfica de la función analizada en el apartado anterior ...

$$p_{dem}(n_{all}, p_1) - p_1$$

...puede describirse como un "gráfico del valor de la ventaja democrática", ya que refleja precisamente el tamaño de la diferencia entre la probabilidad de acertar del colectivo y la probabilidad media de acertar del participante individual.





El gráfico muestra que, independientemente del tamaño del equipo, la mayor ventaja se obtiene en la "zona de confianza débil", donde la probabilidad de acertar es de aproximadamente 0,6. Y, por supuesto, es mayor en la zona de confianza débil. Y, por supuesto, es mayor cuanto mayor es el tamaño del equipo.

Aunque aquí también podría parecer intuitivamente lo contrario, en los casos en los que todo el mundo sabe casi exactamente la respuesta correcta (es decir, la probabilidad media de acertar en todo el colectivo es de

0,95, por ejemplo), la democracia debería mostrar una ventaja especialmente fuerte.

El principal problema de la democracia

Hasta aquí se han considerado los casos en los que la probabilidad de que cada miembro del colectivo acierte es superior a ½. Y en todo este ámbito la democracia tenía ventaja sobre la dictadura.

Sin embargo, lo que determina la ventaja también determina la desventaja.

La probabilidad media de adivinar la opción correcta por parte de los votantes podría ser inferior a ½. Esto es posible, por ejemplo, si hay alguna idea equivocada sobre el estado de cosas que prevalece en el grupo.

En este caso, la probabilidad de seleccionar la respuesta correcta mediante votación bajará a cero con la misma rapidez con la que había bajado a uno.

Al mismo tiempo, el dictador −incluso el que adivina lanzando una moneda− es mucho más eficaz que el equipo, porque adivinará de media una de cada dos veces, mientras que el equipo casi nunca.

Al mismo tiempo, es claramente ventajoso (ajustado a los problemas de igualdad de votos) nombrar dictador a un participante elegido al azar: aunque esa persona adivine incluso peor que la moneda, seguirá adivinando mejor que el equipo.

Ésta es precisamente la razón por la que, en la práctica en particular, puede resultar que un pequeño equipo de especialistas, que tomen las decisiones por votación, sea mucho mejor que un equipo grande, formado principalmente por no especialistas: si en la sociedad existe un delirio sobre una determinada cuestión, aunque esté representado en la misma proporción entre el equipo pequeño, éste seguirá tomando decisiones correctas con mucha más frecuencia (aunque con menos frecuencia que una moneda).

Si tenemos en cuenta que entre el equipo de especialistas el delirio puede estar representado en menor medida, su ganancia será radical: casi una única probabilidad de tomar la decisión correcta frente a casi cero en un equipo grande.

Por último, si el nivel de conocimientos de la sociedad es bajo, el dictador−especialista puede ser uno de los pocos capaces de expresar la opción correcta; al fin y al cabo, al menos debe expresarse para ser seleccionable.

Distribución de probabilidad de la elección correcta

El apartado anterior parecería indicar inequívocamente en contra de la democracia, pero hay una serie de matices con esta evidencia.

La ignorancia puede estar presente, pero no lo está en absoluto siempre: incluso ante un conocimiento no muy bueno de algún área y un pensamiento crítico no muy desarrollado, las personas toman con bastante frecuencia la decisión correcta con una probabilidad mayor que ½ en lugar de menor a la hora de elegir entre dos opciones. Porque claro, hay cosas que no son obvias, hay cosas que requieren conocimientos especiales, hay cosas que son engañosas, pero no todas las cosas son así. Y no para todas las cosas la poca experiencia en temas relacionados con ellas hace que la probabilidad de que la opción equivocada de las dos ofrecidas sea mayor que la correcta.

Al fin y al cabo, con un desconocimiento total de la respuesta correcta a una pregunta, en contraposición a la ilusión, la probabilidad de elegir la respuesta correcta es exactamente ½, no "menos de ½". En consecuencia, para que la probabilidad media de que cada participante elija correctamente sea inferior a ½, no basta con que la mayoría no conozca la respuesta correcta, sino que tiene que "conocer" la incorrecta. Es decir, creer que la respuesta incorrecta es correcta con un grado razonable de certeza. En otras palabras, estar bajo el poder del engaño.

Aunque es más probable que el colectivo, en su mayor parte bajo el imperio de la ilusión, tome una decisión equivocada sobre una cuestión relacionada con esa ilusión que un dictador, sin embargo, deberíamos fijarnos en la situación media de todas las cuestiones a la hora de evaluar la eficacia de la democracia o la dictadura.

Es poco probable que, en general, todas las respuestas correctas a todas las preguntas estén en desacuerdo con alguna de las ideas preconcebidas de la mayoría de los participantes. Es poco probable que incluso la respuesta correcta a todas las demás preguntas las contradiga. Al contrario, el hecho mismo de que la humanidad no sólo no se haya extinguido, sino que incluso siga evolucionando, debería indicarnos que, por término medio, aparentemente las personas tienen más probabilidades de tomar la decisión correcta que la incorrecta a la hora de elegir entre dos opciones.

Por esta razón, si estamos considerando una pregunta sobre la que no sabemos de antemano si hay delirio dominante en ella, no podemos decir inequívocamente que la probabilidad es seguro que sea inferior a ½. Al contrario, tiene sentido suponer que es probable que sea superior a ½; después de todo, en menos de la mitad de las preguntas hay delirios dominantes. Es probable que ni siquiera exista un 10% de ellas.

Del mismo modo, puede haber áreas enteras en las que los prejuicios de un determinado equipo tengan más probabilidades de equivocarse que de acertar, pero es poco probable que estas áreas contengan más de la mitad de las preguntas decididas por ese equipo.

Es decir, aunque en algunas áreas la probabilidad media de elegir la opción correcta sea inferior a ½, la probabilidad media de elegir la opción correcta para todo el conjunto de preguntas sigue siendo superior a ½.

Probabilidad de elegir la opción correcta en todo el conjunto

Intentemos estimar la probabilidad de elegir la opción correcta para aquellas condiciones en las que las probabilidades de elegir la opción correcta son diferentes para las distintas preguntas.

En otras palabras, hay, por ejemplo, algunas preguntas para las que cada participante elegirá por término medio la opción correcta con una probabilidad de 0,60, algunas preguntas para las que la probabilidad de elegir correctamente es de 0,40, otras de 0,90, y así sucesivamente.

Sin embargo, es probable que el número de preguntas con una probabilidad de elección correcta de 0,60 no sea igual al número de preguntas con una probabilidad de elección correcta de 0,40.

La relación cuantitativa entre grupos de preguntas puede describirse mediante una función de densidad de probabilidad, que tiene un significado similar al que suele mostrarse en los histogramas: cuántos elementos se encuentran dentro de un intervalo determinado. La diferencia es que cada "barra" de la función de densidad de probabilidad no contendrá el número de preguntas con esa probabilidad concreta, sino su proporción respecto al total. Y las columnas tendrán un grosor tendente a cero.

Esta función, al mismo tiempo, determina con qué probabilidad una pregunta elegida al azar del conjunto completo de preguntas caerá en uno u otro grupo, es decir, tendrá alguna probabilidad específica de respuesta correcta.

Nótese que ahora hay dos probabilidades en el razonamiento, y el suceso "se elige correctamente", se ha convertido en compuesto.

Es decir, en primer lugar se elige una pregunta al azar y, a continuación, se elige la respuesta correcta para esta pregunta con una determinada probabilidad. La probabilidad de elegir una pregunta con una cierta probabilidad de la respuesta correcta viene dada por la función de densidad de probabilidad.

Así, la probabilidad de la respuesta correcta para cada decisión es

$$P = p_{\text{выбора вопроса}} * p_{\text{правильного ответа на этот вопрос}}$$

Sin embargo, ahora no necesitamos hallar la probabilidad de éxito en una pregunta concreta, sino la probabilidad de éxito en todo el conjunto de preguntas.

Para hallar esta probabilidad, tenemos que resumir todos los trabajos posibles como el anterior.

Te lo explicaré con un ejemplo más sencillo.

Te presentas a un examen. Tienes un total de 10 preguntas para el examen y sabes que 2 de ellas son difíciles y 8 fáciles. Tienes una probabilidad de 1/8 de responder correctamente a las difíciles y una probabilidad de 3/4 de responder correctamente a las fáciles.

¿Cuál es la probabilidad de que te toque una pregunta fácil y la contestes correctamente?

$$\frac{8}{10} * \frac{3}{4}$$

¿Qué probabilidades hay de que te pongan una pregunta difícil y la contestes correctamente?

$$\frac{2}{10} * \frac{1}{8}$$

Pero, ¿cuál es la probabilidad global de acertar la respuesta? Es la suma de la probabilidad de sacar una pregunta sencilla y responder correctamente, más la probabilidad de sacar una pregunta difícil y responder correctamente.

$$\frac{8}{10} * \frac{3}{4} + \frac{2}{10} * \frac{1}{8}$$

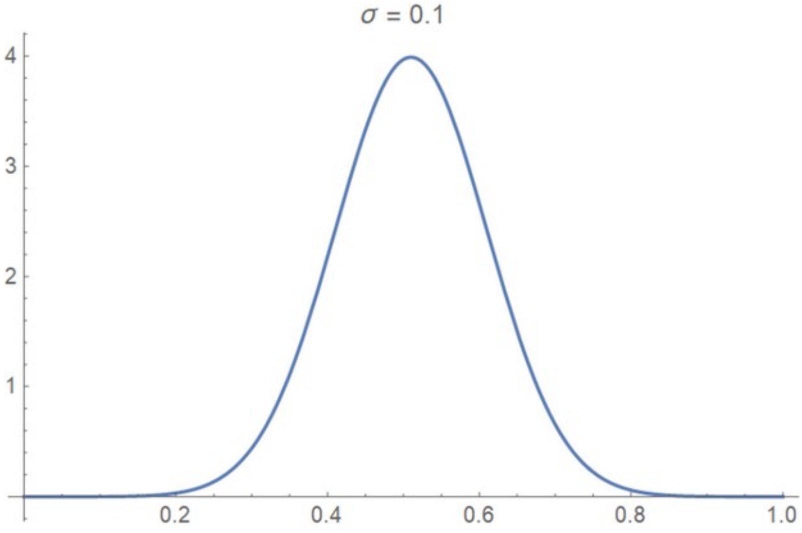
Ahora se buscará un valor similar: la probabilidad de responder correctamente a todo el conjunto.

Se supondrá que la distribución de las preguntas según sus probabilidades de respuesta correcta es gaussiana, es decir, una distribución normal.

La densidad de probabilidad de una distribución normal se define de menos infinito a más infinito.

$$p_{\mu,\sigma}(x) = \frac{\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\,\pi}}$$

donde μ es la media de la distribución y σ es su desviación típica (que determina la anchura de la curva a la mitad de su altura).



Dado que la probabilidad de ocurrencia de cualquier suceso de su conjunto completo es uno, es cierto para esta función

$$\int_{-\infty}^{+\infty} p_{\mu,\sigma}(x)\,\mathrm{d}x = 1$$

Sin embargo, en el caso que nos ocupa, se introduce una distribución normal para las probabilidades de respuesta correcta, que no puede ser inferior a cero ni superior a uno, por lo que éste es el segmento en el que se sitúan absolutamente todas las preguntas. En consecuencia, hay que introducir un factor de normalización de la función de densidad de probabilidad. Será igual a

$$norm_{\mu,\sigma} = \int_0^1 p_{\mu,\sigma}(x)\,\mathrm{d}x$$

En general, la fórmula de este coeficiente es bastante engorrosa.

$$norm_{\mu,\sigma} = \frac{\operatorname{erf}\left(\frac{1-\mu}{\sqrt{2}\,\sigma}\right) + \operatorname{erf}\left(\frac{\mu}{\sqrt{2}\,\sigma}\right)}{2}$$

donde

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$$

Por lo tanto, aparecerá simplemente con su nombre abreviado.

La función de densidad de probabilidad normalizada será por tanto

$$n_{\mu,\sigma}(x) = \frac{p_{\mu,\sigma}(x)}{norm_{\mu,\sigma}}$$

Para hallar la probabilidad de tomar una decisión correcta en todas las preguntas, multiplique la probabilidad de tomar una decisión correcta en alguna pregunta por la probabilidad de que se produzca esa pregunta (que viene dada por la función de densidad de probabilidad) y, a continuación, integre el resultado de cero a uno. Esto es análogo a lo que se hizo en el ejemplo del examen para hallar la probabilidad de aprobarlo.

$$P(\mu, \sigma) = \int_0^1 p(x) * n_{\mu,\sigma}(x)\,\mathrm{d}x$$

En el caso de un dictador es sencillo, la probabilidad de decisión para alguna pregunta coincide con su coordenada X en la distribución de probabilidad de la pregunta.

$$p(x) = x$$

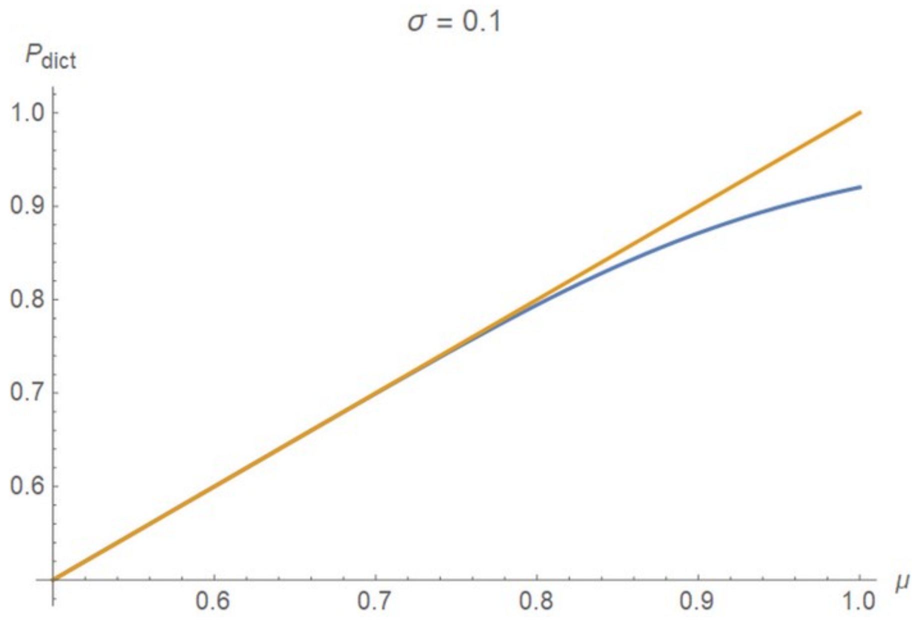
$$P_{dict}(\mu, \sigma) = \int_0^1 x * n_{\mu,\sigma}(x)\,\mathrm{d}x$$

Si te interesa, aquí tienes el aspecto de la fórmula completa con la integral tomada en forma analítica.

$$P\_dict\ (\mu, \sigma) = \sqrt{\frac{2}{\pi}}\sigma \frac{\mathrm{e}^{-\frac{\mu^2}{2\sigma^2}} - \mathrm{e}^{-\frac{(\mu-1)^2}{2\sigma^2}}}{erf\left(\frac{1-\mu}{\sqrt{2}\sigma}\right) + erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)} + \mu$$

Sin embargo, es más útil observar el gráfico que muestra la relación entre la probabilidad de una respuesta correcta en todo el conjunto de preguntas y la probabilidad media de una respuesta correcta para ese conjunto.

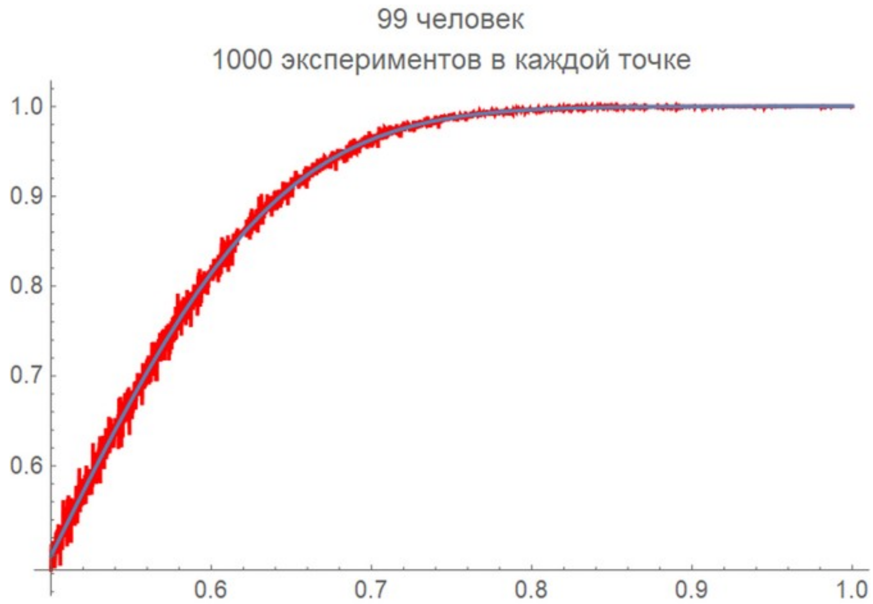
Para mayor claridad, también hay un gráfico de la probabilidad de que un dictador responda a una pregunta concreta con la correspondiente probabilidad de respuesta correcta en el eje X (se trata simplemente de una línea con una pendiente de 45°).



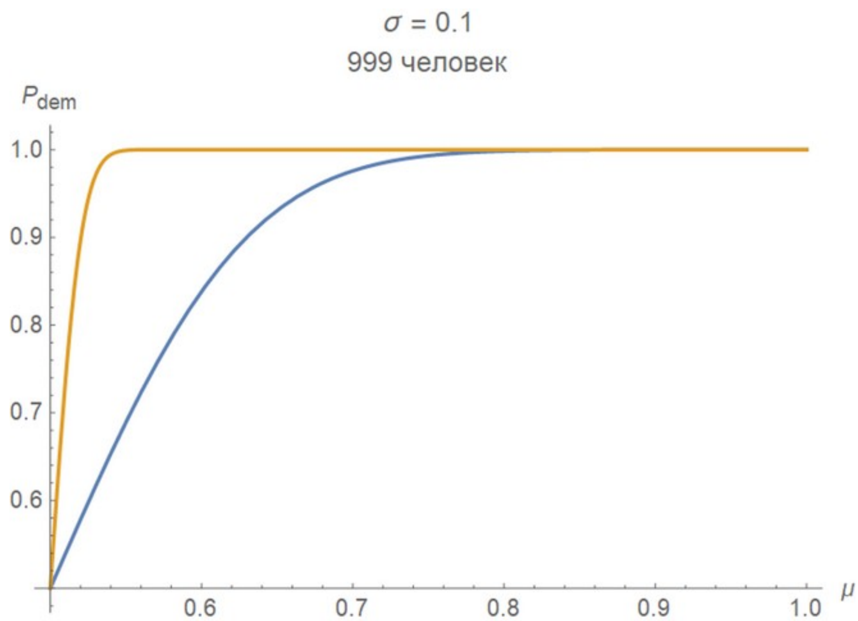
Para la votación democrática, la probabilidad de una decisión correcta viene dada por la función derivada anteriormente

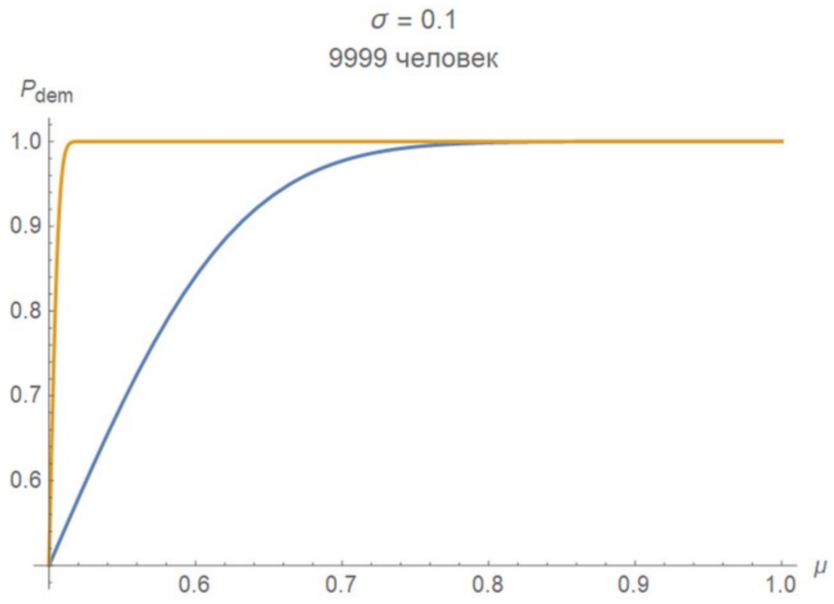
$$p_{dem}(n_{all}, p_1) = \sum_{k=\left\lfloor\frac{n_{all}}{2}\right\rfloor+1}^{n_{all}} \binom{n_{all}}{i} * p_1^k * (1-p_1)^{n_{all}-k}$$

Si hay alguna duda sobre la validez de todas estas construcciones, me gustaría subrayar por separado que esta fórmula también se ha verificado experimentalmente.

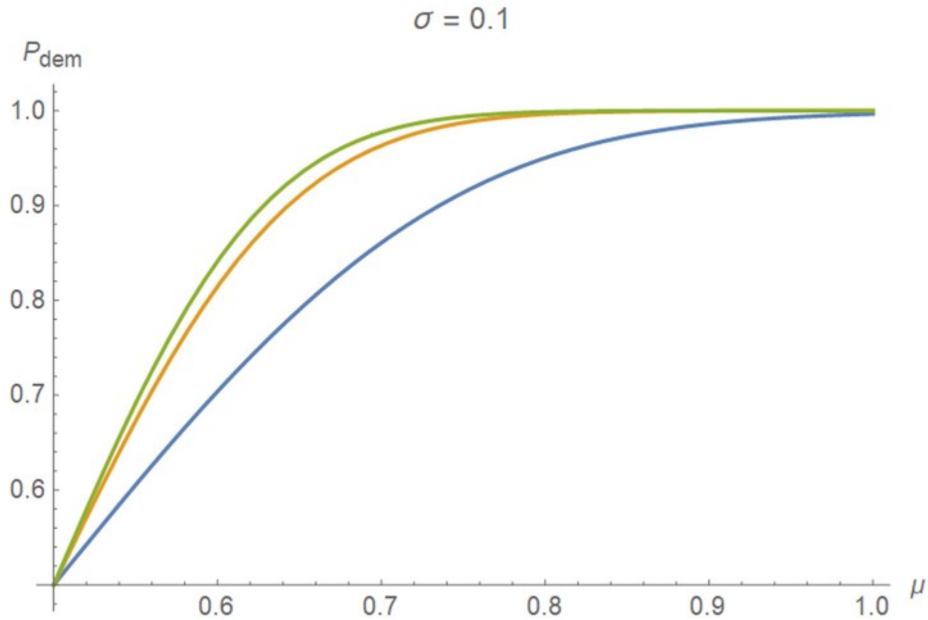


Así es como se ve en los gráficos para algunos números de participantes (aquí también, para comparar, están los gráficos de la probabilidad de que un colectivo responda correctamente a una pregunta concreta con la correspondiente probabilidad coordinada en el eje X).





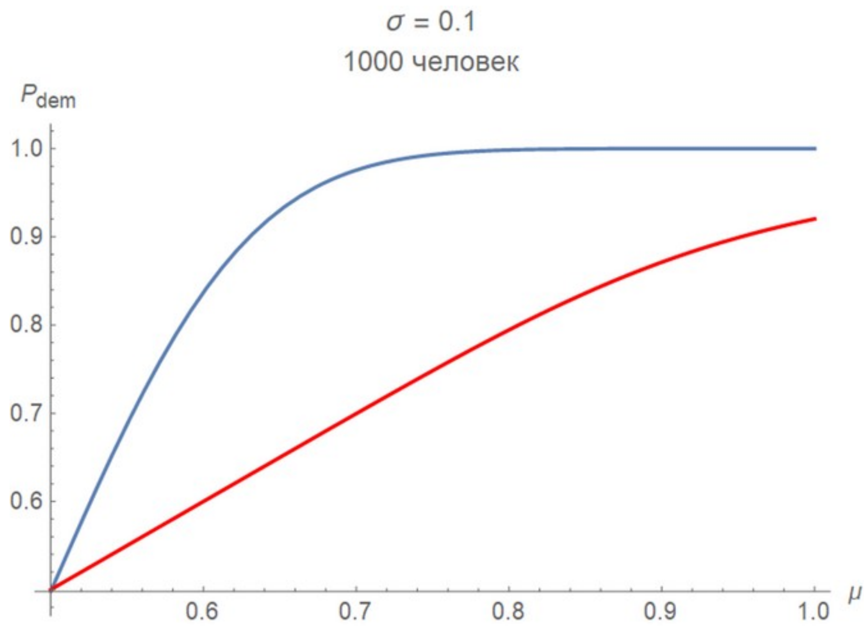
Se puede observar que la probabilidad de hacer una elección correcta en todo el conjunto de preguntas no tiende a uno tan rápido como tiende a aumentar la probabilidad media distribucional de hacer una elección correcta, al igual que la probabilidad de hacer una elección correcta en una pregunta concreta cuando tiende a aumentar la probabilidad de hacer una elección correcta en esa pregunta.



Sin embargo, cuando el número de personas es lo suficientemente grande, la probabilidad de hacer una elección correcta en todo el conjunto se aproxima bastante a uno, aunque la media se desvíe de ½ en todo el conjunto de preguntas.

Así, si la media de la distribución de la probabilidad de una elección correcta en el conjunto de preguntas para 1.000 personas es de 0,74, la probabilidad de una elección correcta en el conjunto será de 0,99, y si la media es de 0,63, la probabilidad de una elección correcta en el conjunto será de 0,9. Esto no es mucho mejor que la probabilidad de una elección correcta en el conjunto de preguntas.

No es mucho mejor que la del dictador para preguntas concretas, pero sigue siendo notablemente mejor.

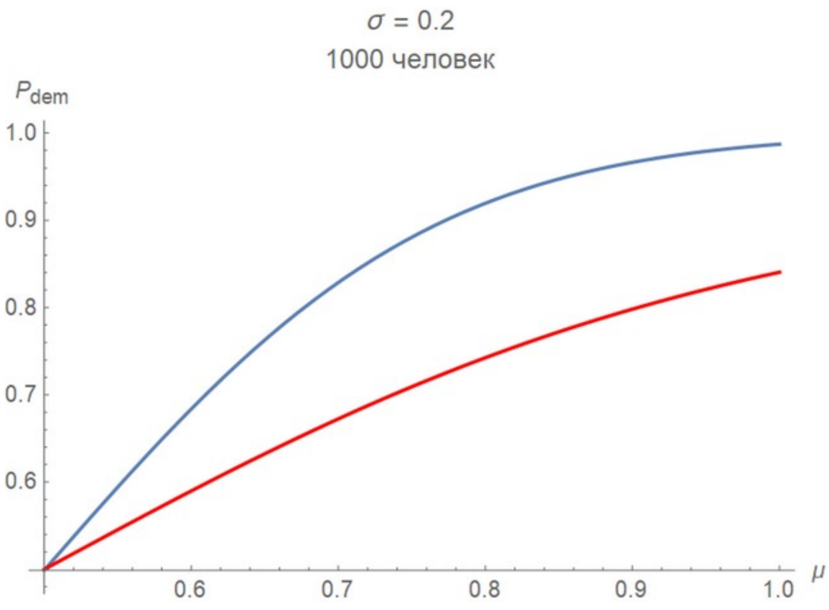


Por cierto, observe que en el conjunto de las preguntas, aunque la probabilidad media de hacer la elección correcta sea uno, la probabilidad de que el dictador haga la elección correcta sigue siendo inferior a uno. En el caso de la democracia, se vuelve casi indistinguible de uno incluso si la probabilidad media sobre el conjunto es 0,75.

Esto determina la existencia de un intervalo de probabilidad promediada sobre el conjunto, en el que la probabilidad de tomar la decisión correcta sobre el conjunto de preguntas votando es tan alta que es inalcanzable para el dictador. Y esto ya no es "prácticamente", sino en principio: si la probabilidad media de una respuesta correcta sobre el conjunto para un colectivo es superior a aproximadamente 0,65, entonces no existe dictador alguno que pueda ser comparable a este colectivo en el mismo valor de la horquilla de probabilidad −para ello la horquilla de probabilidad media tendría que ser superior a la unidad.

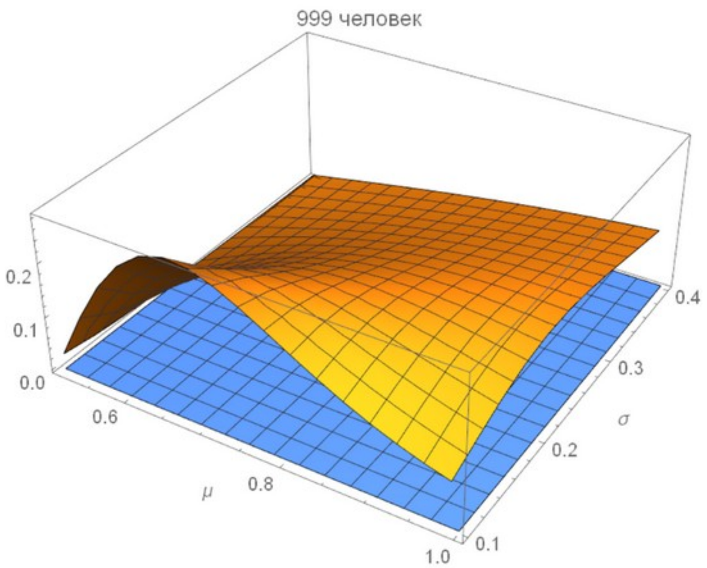
Para que −en caso de que el colectivo se situara en esta franja− "fuera posible", el dictador no sólo tendría que llevar la probabilidad media de responder correctamente a las preguntas en todo su conjunto casi a la unidad, sino también reducir la variación de estas probabilidades. Lo que, en efecto, le lleva a "responder siempre correctamente a todas las preguntas en general".

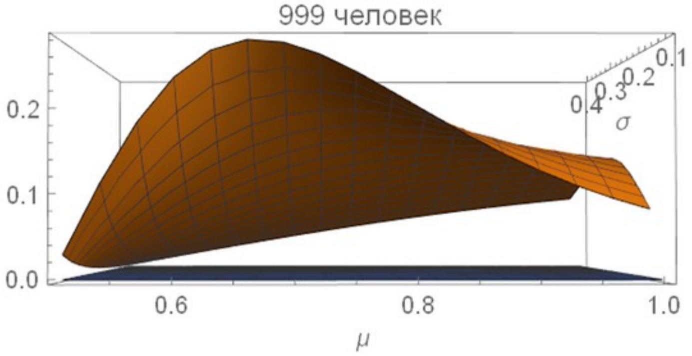
Tomar una desviación estándar mayor, como 0,2, haría que ambos gráficos se desviaran más de la unidad.



Sin embargo, el gráfico "democrático" seguirá siendo notablemente más alto que el "dictatorial". Y aquí también hay una zona de "imposibilidad de un dictador similar" (alrededor de 0,69 en el eje X).

El gráfico de la "ventaja de la democracia" −es decir, la diferencia de probabilidad para el colectivo y para los dictadores bajo los mismos parámetros− en todo el conjunto de preguntas, en función de la media sobre todo el conjunto (μ) y la desviación típica (σ) tiene este aspecto (el plano cero está marcado de nuevo en azul).





Así pues, también en un conjunto completo de cuestiones, la democracia hará que, por término medio, se tomen decisiones correctas con notable mayor frecuencia que en una dictadura.

Sí, el dictador elegirá a veces correctamente donde el colectivo se habría equivocado, pero esto será notablemente más raro que lo contrario. Y esto será así incluso en algunos casos en los que la probabilidad media del dictador de acertar en todas las preguntas sea notablemente superior a la de cada miembro del colectivo.

Por supuesto, es muy posible que el dictador no se equivoque en algo crítico, cuando el colectivo, sujeto al engaño en esta misma cuestión, se equivoque. Pero también puede ocurrir exactamente lo contrario en cuestiones críticas. Y si la media para todas las cuestiones críticas (y no para algún grupo local de ellas) sigue siendo superior a ½ la probabilidad de acertar en el colectivo, entonces el colectivo será más propenso a acertar también en cuestiones críticas.