

---

## Author index

---

Note: with references having more than three authors, the names of all authors appear in this index, but only the first authors appear in the in-text citations.

- Aaronson, D., 231  
 Abayomi, K., 542  
 Achen, C., 194, 484  
 Afshartous, D., 485  
 Agodini, R., 229  
 Agresti, A., 17, 26, 105, 131  
 Ainsley, A., 194  
 Aitkin, M., 484  
 Akaike, H., 105, 527  
 Albert, A., 105  
 Albert, J., 385  
 Aldy, J., 198  
 Almond, D., 231  
 Amemiya, T., 131  
 Anderson, D., 131  
 Anderson, J., 105  
 Angrist, J., 229, 230  
 Ansolabehere, S., 197, 321  
 Ashenfelter, O., 231  
 Asimov, D., 563  
 Atkinson, A., 74
- Bafumi, J., 74, 105, 321  
 Baker, S., 321  
 Balke, A., 231  
 Ball, S., 194, 231  
 Bannerjee, S., 298  
 Barnard, J., 230, 297, 543  
 Barry, S., 342  
 Bates, D., 277, 297, 573  
 Beal, S., 558  
 Beck, N., 248  
 Becker, R., 573  
 Belin, T., 542  
 Benane, S., 485  
 Bentler, P., 485  
 Berger, J., 276  
 Berk, R., 49  
 Berliner, L., 298  
 Berman, B., 297  
 Bernard, H., 342  
 Bernardo, J., 277, 413  
 Bertin, J., 562
- Besag, J., 194, 298  
 Best, N., 105, 298, 527, 573  
 Bird, C., 294, 298  
 Blackman, C., 485  
 Blalock, H., 74  
 Blanchard, S., 527  
 Blattner, F., 132  
 Bloom, H., 229  
 Bock, D., 26, 49  
 Bock, R., 276  
 Bogart, W., 231  
 Bogatz, G., 194, 231  
 Bolger, N., 194  
 Boscardin, W., 297, 434  
 Boshuizen, H., 542  
 Bosker, R., 245, 247, 276, 454  
 Box, G., 74, 276, 413, 454, 526  
 Bradley, R., 132  
 Bradlow, E., 231  
 Brainard, J., 26  
 Brewer, D., 485  
 Bring, J., 74  
 Brookmeyer, R., 194  
 Brooks, S., 342, 373  
 Brooks-Gunn, J., 49, 229  
 Brown, C., 248  
 Browne, W., 297, 485  
 Browner, W., 26  
 Bryk, A., 247, 276, 297, 454  
 Buenconsejo, J., 229  
 Buja, A., 563  
 Burmaster, D., 26  
 Bush, R., 165, 526  
 Butler, M., 230
- Campbell, D., 194, 230  
 Campbell, J., 233, 298  
 Campbell, S., 194  
 Card, D., 231  
 Carlin, B., 105, 276, 298, 373, 413, 434  
 Carlin, J., 26, 49, 105, 151, 165, 248, 276, 298, 346, 385, 386, 413, 434, 526, 574  
 Carroll, R., 74, 230  
 Casella, A., 131, 342  
 Casella, G., 245, 484, 501  
 Catchpole, E., 342

- Chambers, J., 49, 562, 573  
 Chang, H., 526  
 Chapman, R., 131  
 Chay, K., 231  
 Chen, M., 385  
 Chib, S., 385  
 Chipman, H., 298  
 Clarke, B., 485  
 Clayton, D., 342, 542  
 Cleveland, W., 49, 97, 562  
 Clinton, J., 133  
 Clogg, C., 542  
 Cochran, W., 26, 229, 292, 297, 454  
 Cole, N., 276  
 Collins, R., 386  
 Congdon, P., 373, 574  
 Cook, D., 563  
 Cook, R., 8, 105, 485, 526  
 Cook, S., 165, 434  
 Cook, T., 194  
 Costa, D., 198  
 Coull, B., 17, 26  
 Cox, D., 74, 194  
 Cox, G., 197  
 Crainiceanu, C., 230, 298  
 Cramer, C., 105  
 Criqui, M., 74  
 Cromwell, B., 231  
 Cui, Y., 298
- D'Agostino, R., 229  
 Daniels, M., 297  
 Datta, G., 298  
 David, M., 542  
 Davidian, M., 247  
 Dawid, A., 194  
 De Leeuw, J., 8, 245, 247, 248, 276, 297, 485  
 De Veaux, R., 26, 49  
 Dehejia, R., 229, 231  
 Dellaportas, P., 385  
 Dempster, A., 297, 413  
 Dey, D., 526  
 Diamond, A., 229  
 Diffendal, G., 542  
 Diggle, P., 231  
 Dismukes, R., 297  
 Dobson, A., 131  
 Dodhia, R., 563  
 Donohue, J., 321  
 Dorman, P., 198  
 Drake, C., 229  
 Dreze, J., 512  
 Du, J., 230  
 DuMouchel, W., 277  
 Duncan, D., 131
- Dunn, G., 542  
 Dunne, A., 558  
 Dunson, D., 298  
 Dyke, G., 194  
 Dynarski, M., 229
- Earls, F., 342  
 Efron, B., 277, 413  
 Ehrenberg, A., 563  
 Ehrenberg, R., 485  
 Elliott, D., 485  
 Endicott, J., 229  
 Erdos, P., 334  
 Erikson, R., 133, 197, 298  
 Evans, W., 230  
 Everitt, B., 574
- Fagan, J., 131, 342  
 Fair, R., 298  
 Fay, R., 298, 542  
 Felton, J., 51  
 Fienberg, S., 131  
 Finkelstein, M., 543  
 Firth, D., 105  
 Fisman, R., 322  
 Forbes, A., 49  
 Fournier, D., 574  
 Fox, J., 49, 105, 484, 562, 573  
 Frangakis, C., 194, 230  
 Friedlander, D., 229  
 Friedman, J., 132, 298  
 Friedman, M., 132  
 Friendly, M., 563  
 Frolich, M., 229
- Garfinkel, I., 248, 543  
 Gatsonis, C., 277  
 Gawron, V., 297  
 Gelfand, A., 298, 385, 413, 434, 526  
 Gelman, A., 26, 49, 74, 105, 131, 151, 165, 194, 197, 230, 248, 276, 297–299, 321, 342, 346, 373, 385, 413, 434, 484, 501, 526, 542, 543, 563, 574
- Geogebur, Y., 527  
 George, E., 298  
 Gerber, A., 233  
 Geweke, J., 165, 434  
 Gilks, W., 413, 527, 573  
 Gilovich, T., 26  
 Giltinan, D., 247  
 Girosi, F., 298  
 Glickman, M., 230, 277  
 Goegebur, Y., 105  
 Goldstein, H., 247, 276, 454, 485  
 Graddy, K., 230  
 Graubard, B., 485

## AUTHOR INDEX

603

- Graves, T., 574  
Green, B., 245, 501  
Greenland, S., 74, 231, 294, 298  
Greenstone, M., 231  
Guldemon, H., 454  
Gustafson, P., 298, 485
- Haavelmo, T., 230  
Hagstrom, P., 198  
Hahn, J., 230  
Hahn, T., 194  
Hale, R., 294, 298  
Hambleton, R., 321  
Hamermesh, D., 51  
Han, W., 49  
Handcock, M., 342  
Hansen, B., 229, 230  
Hanushek, E., 485  
Harrell, F., 74  
Harris, J., 277  
Hasbun, R., 229  
Hastie, T., 132, 298  
Hauer, E., 131  
Heckman, J., 194, 229, 231  
Heinze, G., 105  
Heisey, D., 454  
Heitjan, D., 153, 542  
Henderson, C., 276  
Herring, A., 298  
Herriot, R., 298  
Higdon, D., 194, 298  
Hill, B., 413  
Hill, C., 229  
Hill, J., 49, 194, 229, 230  
Hills, S., 385  
Hirano, K., 194, 229, 230  
Hodges, J., 298, 385, 434, 527  
Hoenig, J., 454  
Hoeting, J., 298  
Hoff, P., 342  
Honaker, J., 542  
Hoogerheide, L., 512  
Hoppin, J., 298  
House, D., 485  
Hox, J., 276  
Huang, Z., 197, 434  
Hunter, J., 454  
Hunter, W., 454  
Hurley, C., 563
- Ibrahim, J., 385  
Ichimura, H., 194, 229  
Imai, K., 131, 229  
Imbens, G., 194, 229, 230  
Iyengar, S., 322
- Jackman, S., 133, 321  
Jackson, C., 298  
Jackson, P., 276, 342  
Jacob, B., 230  
James, W., 277  
Jaynes, E., 413  
Jeffreys, H., 413  
Jenkyn, J., 194  
Joffe, M., 229  
Johnsen, E., 342  
Johnson, V., 342  
Jones, K., 485  
Joseph, A., 542
- Kahn, M., 198  
Kaldor, J., 342  
Kamenica, E., 322  
Kane, T., 458, 485  
Karagas, M., 230  
Karim, M., 342, 385  
Kashy, D., 194  
Kass, R., 297, 373, 413  
Katz, J., 197, 230, 248  
Kazarow, K., 194  
Kedar, O., 248  
Keller, M., 229  
Kelling, G., 342  
Kempthorne, O., 276  
Kendzioriski, C., 132  
Kenny, D., 194  
Kenward, M., 231  
Kerman, J., 151, 413, 574  
Killworth, P., 342  
King, G., 49, 151, 197, 229, 298, 485, 542  
Kirk, R., 501  
Kiss, A., 131, 321, 342, 484, 543  
Klebanov, P., 229  
Kleibergen, F., 512  
Kleiner, B., 49, 562  
Knook, D., 542  
Korn, E., 485  
Kramer, B., 321  
Krantz, D., 26, 198  
Kreft, I., 8, 245, 247, 248, 276, 297, 485  
Krueger, A., 230, 231  
Kutner, M., 49  
Kwan, E., 563
- Lahiri, P., 298  
LaLonde, R., 229, 231  
LaMotte, L., 245  
Lancaster, T., 230, 373  
Landwehr, J., 105  
Lane, P., 501  
Lange, K., 132  
Lavori, P., 229

- Leamer, E., 194  
 Lechner, M., 229  
 Lee, D., 230, 231  
 Lefgren, L., 230  
 Lenth, R., 454  
 Lessafre, E., 105  
 Levin, B., 543  
 Levine, D., 231  
 Levitt, S., 197  
 Levy, M., 542  
 Lewis, C., 276, 342  
 Lewis, J., 386  
 Leyland, A., 276  
 Liang, K., 131  
 Liaw, F., 229  
 Liebman, J., 131, 321, 543  
 Ligges, U., 574  
 Lin, C., 198  
 Linden, L., 69, 230  
 Lindley, D., 277, 385, 413  
 Linton, O., 229  
 Little, R., 132, 542  
 Little, T., 321  
 Liu, C., 132, 385, 434, 542  
 Lohr, S., 454  
 Loken, E., 321  
 Longford, N., 276, 484  
 Lord, F., 321  
 Louis, T., 276, 277, 413  
 Lovell, J., 131  
 Lu, H., 105  
 Lu, K., 298  
 Lunn, D., 527, 573  
  
 Ma, G., 542  
 Mack, S., 542  
 MacLehose, R., 298  
 Maddala, G., 131, 512  
 Madigan, D., 298  
 Mahboobeh, S., 194  
 Maiti, T., 298  
 Mallows, C., 105, 527  
 Manski, C., 231  
 Martin, A., 321, 574  
 McCabe, G., 49  
 McCarty, C., 342  
 McCarty, N., 321  
 McCullagh, P., 105, 131, 501  
 McCulloch, C., 245, 484, 501  
 McCulloch, R., 230, 297, 298  
 McDonald, G., 76  
 McDonald, J., 563  
 McFadden, D., 132  
 McKeague, I., 543  
 Meng, X., 165, 297, 413, 433, 434, 526, 542  
 Metropolis, N., 385  
  
 Meyers, M., 543  
 Michalopoulos, C., 229  
 Milliff, R., 298  
 Mitchell, J., 51  
 Molenberghs, G., 247, 276  
 Mollie, A., 298  
 Montgomery, D., 454  
 Moore, D., 49  
 Moretti, E., 230  
 Morgan, B., 342  
 Morris, C., 277, 413  
 Moskowitz, A., 153  
 Mosteller, F., 74, 132, 165, 526  
 Murphy, R., 231  
 Murrell, P., 49, 562, 573  
  
 Nachtsheim, C., 49  
 Neal, R., 373  
 Nelder, J., 105, 131, 501  
 Nepomnyaschy, L., 248  
 Nero, A., 277  
 Neter, J., 49  
 Newman, M., 342  
 Newman, T., 26  
 Newton, M., 132  
 Neyman, J., 194  
 Ng, J., 131  
 Nolan, D., 49, 74, 299  
 Normand, S., 230, 277  
 Novick, M., 276, 321, 342  
 Nychka, D., 298  
  
 O'Hara, B., 574  
 O'Keefe, S., 229  
 O'Malley, A., 298  
 Oakes, J., 512  
 Oudshoorn, C., 542  
  
 Palfrey, T., 131, 342  
 Palmer, S., 231  
 Pardoe, I., 49, 105, 133, 485, 526  
 Park, D., 105, 321  
 Parker, A., 51  
 Pasarica, C., 563  
 Pauler, D., 413  
 Pearl, D., 194  
 Pearl, J., 231  
 Peer, J., 297  
 Persico, N., 74  
 Peto, R., 386  
 Pickles, A., 542  
 Pinheiro, J., 277, 297, 573  
 Plummer, M., 574  
 Pollock, M., 485  
 Poole, K., 230, 321  
 Postlewaite, A., 74

## AUTHOR INDEX

605

- Pourahmadi, M., 297  
 Pratt, J., 485  
 Pregibon, D., 105  
 Price, P., 198, 277, 485
- Quagliarello, V., 229  
 Quinn, K., 321, 574
- Rabe-Hesketh, S., 574  
 Racine-Poon, A., 385  
 Raftery, A., 298, 342  
 Raghunathan, T., 542  
 Ramsey, F., 8, 26, 49, 484  
 Rasbash, J., 297, 485  
 Raudenbush, S., 194, 247, 276, 297, 342, 454  
 Raykov, T., 485  
 Reiter, J., 230  
 Renyi, A., 334  
 Richardson, C., 298  
 Richardson, S., 413  
 Richmond, C., 132  
 Ridder, G., 194, 229  
 Ripley, B., 165, 298, 573  
 Ritov, Y., 229  
 Rivers, D., 133  
 Roberts, G., 434  
 Robins, J., 194, 229, 231, 542  
 Robins, P., 229  
 Robinson, G., 245, 248, 276, 485  
 Rockoff, J., 458, 485  
 Rodden, J., 321  
 Romero, D., 133  
 Rosenbaum, P., 194, 229, 231  
 Rosenbluth, A., 385  
 Rosenbluth, M., 385  
 Rosenstone, S., 298  
 Rosenthal, H., 230, 321  
 Rosenthal, R., 485  
 Rosnow, R., 485  
 Ross, C., 75  
 Rotnitzky, A., 229  
 Rubin, D., 26, 105, 151, 165, 170, 194, 229, 230, 276, 297, 298, 346, 373, 385, 413, 434, 485, 526, 542, 574  
 Ruppert, D., 74, 230, 298
- Sacerdote, B., 194  
 Sahu, S., 434  
 Salganik, M., 342, 413  
 Sampson, R., 194, 342  
 Samuhel, M., 542  
 Sargent, D., 298, 434, 527  
 Schafer, D., 26, 49, 484  
 Schafer, J., 8, 542  
 Schemper, M., 105
- Schenker, N., 542  
 Scheve, K., 542  
 Schlessman, J., 74  
 Schultz, B., 542  
 Schwing, R., 76  
 Scott, A., 454  
 Searle, S., 245, 276, 484, 501  
 Sekhon, J., 229  
 Shadish, W., 194  
 Shao, Q., 385  
 Sheiner, L., 558  
 Shelley, G., 342  
 Shen, W., 277  
 Shinharay, S., 526  
 Shively, W., 248  
 Shoemaker, A., 105  
 Shor, B., 105, 321  
 Silver, R., 454  
 Silverman, D., 74  
 Simonson, I., 322  
 Simonton, D., 133  
 Skaug, D., 574  
 Skogan, W., 342  
 Skrondal, A., 574  
 Sleight, P., 386  
 Smith, A., 277, 385, 413  
 Smith, J., 229  
 Smith, T., 454  
 Snedecor, G., 26, 292, 297  
 Snijders, T., 245, 247, 276, 454  
 Snyder, J., 197, 321  
 Sobel, M., 194, 231, 512  
 Solenberger, P., 542  
 Spiegelhalter, D., 105, 413, 527, 542, 573  
 Spitzer, E., 131  
 Staiger, D., 458, 485  
 Stanley, J., 194  
 Stein, C., 277  
 Stern, H., 26, 105, 151, 165, 276, 298, 346, 385, 413, 434, 526, 542, 574  
 Sternberg, S., 527  
 Stewart, C., 197  
 Stigler, S., 49, 74, 132, 159, 277  
 Stinson, M., 51  
 Stone, M., 526  
 Strathdee, S., 194  
 Stuart, E., 229  
 Sturtz, S., 574  
 Subramanian, S., 485  
 Summers, A., 485  
 Swayne, D., 563
- Tan, W., 413  
 Tanner, M., 385  
 Taylor, J., 132  
 Teller, A., 385

- Teller, E., 385  
 Terry, M., 132  
 Thayer, D., 276  
 Thistlethwaite, D., 230  
 Thomas, A., 527, 573, 574  
 Thomas, N., 229  
 Thurstone, L., 132  
 Tiao, G., 276, 413  
 Tibshirani, R., 132, 298  
 Tinbergen, J., 230  
 Tobin, J., 132  
 Todd, P., 194, 229, 230  
 Tomz, M., 485  
 Tosteson, T., 230  
 Trevisani, M., 105  
 Triest, R., 542  
 Trochim, W., 454  
 Troxel, A., 542  
 Tsai, W., 543  
 Tsui, K., 132  
 Tsutakawa, R., 297  
 Tu, W., 230  
 Tuerlinckx, F., 105, 485, 527  
 Tufte, E., 562  
 Tukey, J., 74, 132, 165, 245, 501, 563  
 Tukey, P., 49, 562  
 Tustakawa, R., 413  
 Tversky, A., 26  
  
 Vaida, F., 527  
 Vallone, R., 26  
 Van Buuren, S., 542  
 van der Klaauw, W., 230  
 van der Linde, A., 105  
 van der Linden, W., 321  
 van Dijk, H., 512  
 van Dyk, D., 131, 229, 434  
 van Geen, A., 105  
 Van Hoewyk, J., 542  
 Van Mechelen, I., 105, 527  
 Varadhan, R., 194  
 Vehovar, V., 542  
 Velleman, P., 26, 49  
 Venables, W., 298, 573  
 Verbeke, G., 247, 276  
 Versteeg, R., 105  
 Vikram, H., 229  
 Viscusi, W., 198  
 Vlahov, D., 194  
 Volinsky, C., 298  
 Von Krosigk, C., 276  
 von Mises, R., 27  
  
 Wahba, S., 229, 231  
 Wainer, H., 231, 321, 485, 562  
 Wakefield, J., 413  
  
 Waldfoegel, J., 49, 229  
 Walker, S., 131  
 Waller, L., 298  
 Wallis, W., 132  
 Wand, M., 298, 563  
 Wang, N., 543  
 Wang, X., 485  
 Warnes, G., 574  
 Wasserman, L., 413  
 Wasserman, W., 49  
 Wedderburn, R., 131  
 Weidner, R., 485  
 Weisberg, S., 8, 49, 526  
 West, M., 298  
 West, V., 131, 321, 543  
 Whang, W., 153  
 Wherry, R., 49  
 Wideman, L., 542  
 Wiens, B., 132  
 Wikle, C., 298  
 Wilkinson, L., 563  
 Wilks, A., 573  
 Wilson, J., 342  
 Witte, J., 294, 298  
 Wittenberg, J., 485  
 Wleziën, C., 298  
 Wolfe, B., 485  
 Wolfe, R., 248  
 Wolfers, J., 321  
 Wolfram, C., 197  
 Wong, W., 385  
 Woolridge, J., 49, 132, 230  
 Wright, G., 321  
 Wu, Y., 434  
  
 Xia, H., 298  
 Xiaofeng, L., 454  
 Xu, R., 485  
  
 Yates, F., 501  
 York, J., 298  
 Yusuf, S., 386  
  
 Zanutto, E., 230  
 Zaslavsky, A., 277, 298, 413, 433, 542  
 Zeger, S., 342, 385  
 Zellner, A., 131  
 Zeng, L., 229  
 Zhao, L., 229  
 Zheng, T., 342, 413  
 Zheng, Y., 105  
 Zhou, A., 230  
 Zhou, X., 230  
 Zimmerman, P., 231  
 Zivot, E., 512  
 Zorn, C., 105

---

## Subject index

---

- 1.6, factor for converting from logit to  
probit, 86, 118, 131
- 100,000 iterations, why we do not  
recommend, 369
- 2 standard deviations, scaling by, 56–57
- 2.8 standard errors for power calculation,  
441
- 5 ways to write the same model, 262–265
- 55,000 residents desperately need your  
help, 23–26
- 6 quick tips, 547–549
- 8 schools, 430–434
- ability parameter, in item-response model,  
315–320
- abline(), 353, 520
- Academy awards, 133
- accessing the simulations from a Bugs fit,  
358
- AD Model Builder, multilevel modeling in,  
572–573
- adding a predictor can increase the  
residual variance, 480–481
- additivity, 46
- adjusted parameters, 382, 420, 422  
compared to raw parameters, 423
- adjusted R, 475
- Advanced Placement exams, 463–464
- age guessing, 299
- age, as regression input, 66, 67
- Akaike information criterion (AIC), 105,  
525–527
- aliasing, a form of nonidentifiability, 318,  
321
- American Indians, 340
- analysis of covariance, 496–498
- analysis of variance (ANOVA), 487–502  
additive data decomposition, 488  
balance, 489  
classical, 487–490  
contrasts, 462–466, 496–498  
degrees of freedom, 488, 499  
F-test, 489  
generalized linear models, 491  
graphical display, 492–495, 497, 499,  
500  
mean square, 489  
method of summarizing a multilevel  
model, 492–494  
model comparison and, 489  
multilevel models and, 490–502  
nested designs, 496  
sum of squares, 488  
unbalanced designs, 496
- analytic calculations and simulation, 148
- Annenberg Election Survey, 311
- apply(), 44, 353, 477
- array(), 308
- arsenic, *see* wells
- as.bugs(), 413
- as.bugs.array(), 400
- as.vector(), 348
- attach.bugs(), 305, 352, 358
- augmented data  
multilevel model, 397  
regression with prior information, 393
- autoregression, 163–165
- available-case analysis, 532
- average causal effect, definition, 173
- average predictive comparison, *see*  
predictive comparison
- axes of graphs, 553
- B-K plot, 312, 313, 321
- balance, *see* causal inference
- Bangladesh, *see* wells
- basement, 254
- Bayesian inference, 142, 276, 345–348,  
392–414  
calibration, 433  
censored data, 405–408  
classical regression, 392  
debugging, 434  
exchangeability, 347  
Gibbs sampler, 397–402  
informal, 143  
informative prior distribution,  
392–393, 427–434  
Metropolis algorithm, 408–409  
multilevel model, 393–402, 413  
picture, 395, 396  
prior distribution, 427–434  
variance parameters, 432–434,  
499–501

- separation in logistic regression, 104
- social networks model, 409–413
- beauty and teaching evaluations, 51
- before-after graph, 555
- behavioral learning experiment, 515–524
- beta-binomial model, 321
- binary data and count data, 117
- binary outcomes, 79
- binned residual plot, 97–101, 105, 559
- binomial distribution, 16
  - bounding trick in Bugs, 382
  - compared to Poisson, 112
- binomial-normal model, 321
- births, 27, 137–139
- black box, 190
- bounding parameters in Bugs, 383
- bounding trick for binomial model in Bugs, 382
- brlr package, 104
- broken windows, 325
- Bugs, 11, 305, 345–386, 567, 573
  - bounding parameters, 383
  - bounding trick for binomial model, 382
  - brute force, why we do not recommend, 369
  - calling from R, 350–352
  - censored data, 405–406
  - classical regression, 360
  - compared to lmer(), 386
  - compared to R, 399–402
  - compiling, 416
  - complete-pooling model, 360
  - confidence-building, 415–417
  - current problems with, 417
  - data, 350, 356, 416
  - data, parameters, and derived quantities, 366–369
  - dbin(), 381
  - dcat(), 383
  - debugging, 415–417
  - derived quantities, 367
  - detailed example, 348–359
  - dgamma(), 430
  - dlnorm(), 383
  - dlogis(), 384
  - dmnorm(), 376
  - dnorm(), 353, 354
  - dt(), 372, 384
  - dunif(), 355
  - dwish(), 377
  - each object can be defined at most once, 368
  - equals(), 384, 405
  - example of graphical output, 351
  - example of slow convergence, 421
  - example of tabular output, 351
  - examples, 353
  - fake-data simulation, 363–365
  - finding errors in a model, 373
  - general principles, 366–369
  - group-level model, 354
  - help, 353, 373, 565
  - I(), 383, 384, 405
  - individual-level model, 353
  - initial values, 350, 356, 369–370, 416
  - inprod(), 361, 378, 379
  - interpreting inferences from fitted model, 352
  - inverse(), 377
  - inverse-variance parameterization, 354, 368
  - knowing when to give up, 419
  - latent data for generalized linear models, 384
  - linear transformation to speed convergence, 419–427
  - logistic regression, 381–382
    - latent-data formulation, 384
  - logit(), 381
  - looping indexes, 353, 367
  - max(), 381
  - mean(), 382
  - min(), 381
  - missing data, 367, 529
  - model as specifications, not instructions, 368
  - model too big for, 336
  - modeled data and parameters, 366, 367
  - multilevel model, 375–386
    - logistic, 381–382
    - non-nested, 380–382
    - Poisson, 382
    - varying intercepts and slopes, 375–379
  - multiple group-level predictors, 379–380
  - multiple varying coefficients, 378
  - n.thin, 518
  - nested subscripts, 372
  - no overwriting in model, 368
  - no-pooling model, 360
  - non-nested model, 380–381
  - noninformative prior distribution, 355
  - nonlinear and nonadditive models, 371
  - number of iterations and sequences, 356–358, 369
  - open-ended modeling, 370–372
  - ordered categorical regression, 383
  - overdispersed Poisson regression, 382



## SUBJECT INDEX

609

- parameters, 350, 356
- postprocessing inferences in R, 359
- pow(), 355
- practical advice, 357, 369–370, 415–434, 547–549
- prediction, 361–363
- prior distribution, 354
- $\hat{R}$ , 358, 369
- raw (unscaled) parameters, 377
- redundant parameters, 419–427
- replicated data for predictive checking, 518
- robit regression, 384
- scaled inverse-Wishart model, 376–380
- sd(), 460
- simple multilevel model, 350
- slow with large datasets, 418
- stochastic learning model
  - logarithmic link, 521
  - logistic link, 517
  - multilevel, 522–524
- thinning, 418
- too slow for looping in power calculation, 451
- unequal variances, 371–372
- unmodeled data and parameters, 366, 367, 378
- varying-intercept model, 350, 361
- varying-intercept, varying-slope model, 375–379
  - group-level predictors, 379–380
  - modeling the correlation, 376
  - no correlation, 375
  - scaled inverse-Wishart, 376
  - why you should learn Bugs, 345
- Wishart model, 376–380
- bugs(), 350, 565, 567
- building generalized linear models, 125–127
- building regression models, 68–73
- c(), 350
- calibration
  - Bayesian inference, 433
  - plot, 309, 558
- calling Bugs from R, 350–352
- Caltech, 332
- cancer and nutrients, 294
- categorical variables as regression inputs, 66
- Cauchy distribution, 428, 430
- causal inference, 95, 167–198, 549
  - average causal effect, 173, 180, 205
  - balance, 181, 184–186, 199–204, 208, 209, 212
  - plot, 202, 229
  - causes of an effect, 187
  - confounding covariate, 169, 176, 181, 184, 196, 200, 202–203, 207, 212–213, 215
  - constructed observational study, 210, 230, 231
  - continuous treatment, 177
  - controlling for post-treatment variable, don't do it, 177, 187–194, 227, 229
  - counterfactual, 170, 173, 184, 185, 188, 201, 206, 218
  - definition, 170
  - difference-in-differences, 228, 231
  - direct and indirect effects, 191–194
  - effect of causes, 187
  - effect of the treatment on the controls and treated, 205, 206, 211
  - exclusion restriction, 216–218
  - experiment, *see* randomized experiment
  - external validity, 174
  - fixed effects, 231
  - fundamental problem, 171–172, 191
  - hypothetical intervention, 187
  - ignorability, 182–184, 186, 207, 212, 216–217, 225, 228, 231, 232
  - infinite regress of causation, 187
  - instrumental variables, 215–226, 230–231, 233
    - assumptions, 216–218, 220–221, 225–226, 233
    - derivation, 218–221
    - identifiability, 220–221, 224
    - multilevel model, 509–511
    - standard error, 222
    - two-stage least squares, 221
    - Wald estimate, 219, 221
    - weak instrument, 225
  - intent-to-treat (ITT) effect, 216, 219, 220, 222
  - internal validity, 174
  - local average treatment effect (LATE), 219–220, 229, 233
  - mapping your study to an ideal randomized experiment, 187–188
  - matching, 206–212
    - model checking, 208
    - nearest-neighbor, 207
    - propensity score, 230
    - R packages for, 230
  - model extrapolation, 169, 185, 201, 209, 213
  - monotonicity, 216–217

- multilevel model, 178, 210, 212, 227–228, 503–512
- multiple treatments, 177, 186, 187
- natural experiment, 187, 225
- nonignorable assignment mechanism, 183
- observational study, 172, 181–188
  - design of, 437
- overlap, 184–186, 199–204, 209, 212–215
- poststratification, 178, 181, 206
- potential outcome, 168, 170, 171, 183, 186, 189, 191, 219
  - close substitutes, 171–172
- predictive comparison, 167, 168
- principal stratification, 192
- propensity score matching, 207–210, 230, 232
  - assumptions, 207
  - model checking, 208
  - R packages for, 230
  - standard errors, 212
- propensity score weighting, 211, 229
- randomized encouragement, 216
- randomized experiment, 171–181, 183
  - completely randomized, 175, 182, 183
  - design of, 437
  - paired comparisons, 175
  - randomized blocks, 175
- regression, 167–198
- regression discontinuity, 212–215, 232
- selection bias, 168
- selection model, 231
- selection on observables, 183
- sensitivity analysis, 231
- simultaneous causality, a meaningless concept?, 188
- structural equation model, 226
- subclassification, 204–207, 210–211, 229
- support of the data, 169
- SUTVA, stable unit treatment value assumption, 178
- treatment interaction, 178–180, 189, 205, 214–215
- treatment variable, definition, 186
- weighting, 205
- `cbind()`, 43, 146, 157, 361, 529
- CD4, 249, 277, 373
- censored data, 402–408
  - Bayesian inference, 405–408
  - Bugs, 405–406
  - Gibbs sampler, 406–408
  - imputation for, 540–541
  - likelihood, 404
  - maximum likelihood, 404–405
  - naive regression estimates, 403–404
  - picture, 541
  - R, 404–408
- Census, 277, 301, 308
- centering, 55–57, 419, 464–466
  - changes the correlation of group-level intercepts and slopes, 288–289
  - interactions and, 93
  - non-nested model, 292
- Central Limit Theorem, 13–14, 27
  - regression coefficients, 14
- chick brains, magnetic fields and, 481–484
- child care, 189–193, 201, 224
- child support enforcement in cities, 237–241
- children with HIV, 249, 443–447
- children's test scores, 31–51, 55–57
- classical models for regression coefficients, 293
- classical regression, as a special case of multilevel modeling, 258
- cluster sampling, 437, 447–449
- clustered data, 237–241
- cockroaches, 126–127, 161–163
- `coef()`, 43, 156, 260, 267, 280, 352
- coin flipping, patterns in, 554
- collinearity, 68, 255, 393, *see also* identifiability
- `colMeans()`, 520
- colochos, canasta de, vii, xxii
- `colors()`, 43
- `colSums()`, 411
- combining inferences from multiple imputations, 542
- combining separate local regressions, 263
- comparing two surveys, 442
- comparisons, graphs as, 552–553
- compiling Bugs, 416
- complete pooling, 247, 252–259, 270, 348
  - Bugs model, 360
  - problems, 256
  - special case of multilevel modeling, 258
- complete-case analysis, 531
- compound models
  - imputing a variable that can be positive or zero, 537–538
  - simulation of, 150–151, 537–538
- computer display, 565, 566
- computing, 9–11, 547, 565–574
  - debugging, 415–417
  - instead of matrix algebra, 8
  - multilevel models, 345–434
  - practical advice, 415–434
- confidence intervals, 18–20

## SUBJECT INDEX

611

- continuous data, 18
- discrete data, 18
- proportions, 18
- confidence-building, 415–417
- confounding covariate, 169, 176, 181, 184, 196, 200, 202–203, 207, 212–213, 215
- congressional elections, 76, 144–148, 197, 213, 233
- conjunctive item-response or ideal-point model, 319
- connect times on the web, 492–493
- Connecticut, what's the matter with, 310–314
- constant term, 251, 349
- constraining a batch of coefficients to sum to zero, 326
- constructed observational study, 210
- constructive choice models, 127–131
- contextual effect, 481
- continuous and discrete predictors, 66
- continuous probability simulation, 152
- contrasts, 462–466
  - analysis of variance, 496–498
  - computing in R, 464
- convergence of iterative simulations
  - monitoring, 352
  - pictures, 357
- correlation, 57–59, 265, 389
  - coefficient estimates, 40
  - graph, 340
  - group-level intercepts and slopes, 279, 287–289
  - individual-level variables and
    - group-level errors, 481
    - modeling in Bugs, 376–380
    - scaled inverse-Wishart model, 287
- cost-benefit analysis, 128, 153, 454–455
- costs and benefits of multilevel modeling, 246–247
- count data and binary data, 117
- counterfactual, 170, 173, 184, 185, 188, 201, 206, *see also* potential outcome
- counterfactual and predictive
  - interpretations of regression, 34
- covariance, *see* correlation
- coverage of confidence intervals, 156
- cows, 196
- $C_p$ , 527
- curve(), 43, 353
- cutpoints, for ordered logit or probit, 119–120, 332
  - estimating in Bugs, 383
- data cleaning, for social networks survey, 333
- data for examples, 11
- data matrix, 37, 238, 239
  - group-level, 239, 240, 243
  - imputation, 541
  - individual-level, 239, 243
  - non-nested data, 243
- data model, 347
- data reduction, 209
- data sent to Bugs, 350, 356, 416
- data subsetting, 326, 357, 547
  - speeding computation, 418
- data.frame(), 48, 140, 452, 535
- dbin(), 381
- dcat(), 383
- death penalty, 19, 116, 243–244, 320–321, 540–541
- debugging, 415–417, 434
  - diagram of general strategy, 416
- decision analysis, well-switching example, 127–131
- default line in graph, 556, 560
- defaults in R functions, 452
- degree distribution, estimated for men and women, 337
- degrees of freedom, 41, 372, 488
  - inverse-Wishart distribution, 286
- Democrats and Republicans, 310–314
- dependent variable, *see* outcome
- derived quantities, 366, 367
- design of sampling and experiments, 437–455
- details in graphs, 560
- deterministic or random imputation, 534, 537
- deterministic part of a regression model, 31
- deviance, 100, 105, 113, 524–526
- deviance information criterion (DIC), 105, 525–527
  - instability in computation of, 526
- dgamma(), 430
- diagnostics
  - external validation, 48
  - residuals, 40, 47, 97–101
  - simulation-based, 155–166, 513–524
- diarrhea, zinc, and HIV, 443–447
- difference-in-differences estimation, 228, 231
- difficulty parameter, in item-response model, 315–320
- dim(), 147
- dimnames(), 400
- discrepancy variable, 513
- discrete and continuous predictors, 66, 107
- discrete probability simulation, 152

- discrimination parameter, in item-response model, 316
- disjunctive item, 319
- display(), 38–39, 565
- distance to the nearest safe well, 88
- distribution, 13–16
- binomial, 16
    - bounding trick in Bugs, 382
  - Cauchy, 428, 430
  - folded noncentral t, 428
  - gamma, 335
  - log=TRUE option in R, 405
  - logistic, 85
    - Bugs, 384
  - lognormal, 15, 383
  - multivariate normal, 15
  - negative binomial, 115, 336
  - normal, 13–15, 263
    - computing in R, 405
    - truncated, 407
  - Poisson, 16, 110–116, 335
  - scaled inverse-Wishart, 284–287, 376–380
  - t, 124, 372, 428
  - truncated normal, 407
  - Wishart, 284–287, 298, 376–380
- divide by 4 rule, for interpreting logistic regression, 82
- dividing by two standard deviations, 56–57
- dlnorm(), 383
- dlogis(), 384
- dmnorm(), 376
- dnorm(), 353, 354, 404
- dog experiment, 515–524
  - model comparison, 526
  - observed and replicated data, 516, 523
- Douglas, William, 318
- dt(), 372, 384
- dummy variable, *see* indicator
- dunif(), 355
- dwish(), 377
- dynamic graphics, 563
- earnings
  - height and, 50, 53–54, 75, 126, 287, 290–292
  - logarithmic models, 59–65
  - mixed discrete/continuous data, 126
- econometrics and biostatistics, 231
- education as categorical input variable, 95
- educational children's television programs, *see* Electric Company *and* Sesame Street
- educational testing, 317, 430–434
- effect size, why more important than sample size, 439
- effective sample size, 352
- elasticity, 64, 76
- election forecasting, graphs of, 557, 562
- election fraud, 23–26
- Electric Company experiment, 174–186
  - graph of data, 552
  - multilevel model, 503–505
- Emacs, 565
- equal variance of errors, 46
- equals(), 384, 405
- Erdos-Renyi model, 334
- error rate, 99
- errors, distinguished from residuals, 387
- exam scores, actual vs. guessed, 558, 559
- examples, data for, 11
- exchangeability and prior distribution, 347
- exclusion restriction, *see* causal inference, instrumental variables, assumptions
- expected(), 123
- experiment, *see* causal inference, randomized experiment
- experimental design, 437–455
- experimental economics, 331
- explained variance, *see*  $R^2$
- exploratory data analysis, 551
- exposure, in Poisson model, 111–113
- expression(), 148, 353
- external validation, 48
  - election polls, 309
- external validity, 174
- F-test, 489
- factor, 68
- factor analysis, 296
- factor(), 255, 349
- fake-data simulation, 50, 155–158
  - Bayesian inference and, 434
  - checking coverage of intervals, 365
  - multilevel model, 363–365
  - residual plots, 157–158
  - sample size and power calculations, 449–454
  - using Bugs, 363–365
- feeling thermometer, 86
- figure skaters, 248
- finite-population variance, 459–462, 499
  - analysis of variance (ANOVA), 491
  - Bugs, 460
  - graph, 461
  - unmodeled coefficients, 462
- fit many models, 547
- fitted(), 158
- fitting multilevel models, 259–262, 345–434
- fitting the wrong model, 165
- fixed effects, 231, 244–246, 259, 261

## SUBJECT INDEX

613

- finite-population inferences and, 461
  - many definitions, 245, 248
  - R, 260
  - why we avoid the term, 2, 226–228, 245
- fixef(), 260, 280
- flight-simulator experiment, 289–290, 297, 464–466
  - analysis of variance, 488
  - Bugs model, 380
  - superpopulation and finite-population variance, 459–460
- folded noncentral t distribution, 428
- forecasting elections, 294–296
- Fragile Families study, 238
- function(), 350, 401
- functions in R, 139, 147, 151, 404, 534, 535
- fundamental problem of causal inference, 171–172, 191
  
- gain scores, 177, 195
- Gallup Report, 560
- GAMM package, 567
- gamma distribution, 335
- generalized additive model, 298, 567
- generalized estimating equations, 248
- generalized linear model, 109–133
  - analysis of variance (ANOVA), 491, 493–494
  - binomial, 116–117
  - building, 125–127
  - cutpoints, 119–120
  - deviance, 100, 105, 113
  - latent-data formulation, 384
  - logistic, 79–108
  - logistic-binomial, 109, 116–117
  - multilevel, 325–342
  - multinomial logit and probit, 110, 119–124
  - negative binomial, 115
  - ordered logit, 119–124
    - Bugs, 383
    - multilevel, 331–332
  - others, 127
  - Poisson, 109–116
    - Bugs, 382
    - compared to binomial, 112
    - exposure, 111–113
    - overdispersion, 382
  - prediction, 272
  - predictive comparisons, 466–473
  - probit, 109, 118–119
  - probit or logit, 118
  - robit as robust alternative to logit and probit, 124–125
  - robust, 110
  - simulation, 148–151
  - thresholds, 119–120
- Gibbs sampler, 385, 397–402
  - censored data, 406–408
  - linear transformation to speed
    - convergence, 419–427
  - model building and, 402
  - multilevel model, 398–402
  - picture, 398
  - programming in R, 399–402, 412
  - redundant parameters, 419–427
  - slow convergence, 424
  - social networks model, 410
  - updating functions in R, 399–402
- glm(), 79, 110, 565
- global variables in R, 400, 412
- goodness-of-fit, *see* model checking
- grades, predicting, 157
- graphics, 42–45, 548, 551–563
  - comparisons, 552–553
  - general advice, 562
  - instead of tables, 328, 337
  - jittering, 32
  - no single graph does it all, 552–553
  - plotting regression coefficients, 337, 341
  - plotting symbols, 555
  - R, 562
  - scatterplot with regression lines
    - superimposed, 35, 42–45
  - shape of the plotting region, 556
  - showing fitted models, 551, 553
  - symbols and auxiliary lines, 554
  - theory of, 562
  - why, 551
- group, 261
- group indicators, 264
- group- and individual-level data matrices, 239, 243
  - imputation, 541
- group-level predictors, 265–269, 271
  - along with group indicators, 269, 293, 463–464, 498
  - Bugs model, 361
  - varying-intercept, varying-slope models, 280
- group-level standard deviation, 270
- guessing ages, 299
- guessing, in item-response model, 319
  
- handedness, 66
- hard constraint, 257
- hazard regression, 298
- height, 139
  - earnings and, 50, 53–54, 75, 126, 287, 290–292

- logarithmic models, 59–65
  - mixture model for, 14
  - parents and children, 58
  - weight and, 41, 74, 402–408
- help in Bugs, 565
- help in R, 405, 565, 567
- heteroscedasticity (unequal variances), 297
- hett package for robust regression, 110, 124, 133, 567
- hist(), 137, 536, 562
- histograms, 561
- HLM, multilevel modeling in, 573
- holding other predictors constant,
  - difficulty of, 34
- homeless people, 333
- hot-deck imputation, 538
- how many groups, 275–276
- how many observations per group, 275–276
- how many x's survey, 332–342
- hyperparameter, 1, 258
- hypothesis testing, 20–26
  
- I(), 215, 383, 384, 405, 538
- ideal-point model, 314–321
  - multilevel, 316
  - picture, 315
  - redundant parameters in Bugs, 426
  - two-dimensional, 319
- identifiability, 419, 420
  - Bayesian regression and, 393
  - categorized predictors and, 68
  - causal inference, 170
  - constant term in non-nested models, 381
  - ideal-point model, 318
  - instrumental variables, 220–221, 224
  - item-response model, 315
  - likelihood and, 392
  - linear regression, 68
  - logistic regression, 86, 104, 107
  - social networks model, 336
- ifelse(), 126, 150, 384, 403, 534
- ignorability, 182–184, 186, 231, 530, 542,
  - see also* causal inference
- imbalance, *see* causal inference, balance
- imputation, *see* missing data
- impute(), 535
- income and voting, 79–84, 105, 107, 310–314
- incremental cost-effectiveness ratio, 153
- incumbency, 197, 233
- independence of errors, 46
- independent variable, *see* input variable
- index number, don't graph by, 553
- index variable, 67, 238, 252
  - creating in R, 348
  - non-nested models, 289
- indicator variable, 67, 238, 244–246, 255
  - default, reference, or baseline condition, 68
- individual- and group-level data matrices, 239, 243
  - imputation, 541
- Infant Health and Development Program,
  - see* child care
- inference, *see* statistical inference
- informative prior distribution, 392–393
- initial values sent to Bugs, 350, 356, 369–370, 416
  - restricted parameter space, 384
- inprod(), 361, 378, 379
- input variables, as distinguished from predictors, 37, 466
- install.packages(), 567
- instrumental variables, *see* causal inference
- interactions, 34–36, 242, 453
  - centering the input variables, 93
  - graphing, 36, 94, 313
  - logistic regression, 92–96
  - predictive comparisons and, 469
  - sample size and, 438
  - treatment effects, 178–180, 189, 205
  - varying slopes as example of, 282–283
  - when to look for, 36
- intercept, 33, 35, *see* varying intercepts and slopes
- intermediate outcome, *see* causal inference, controlling for post-treatment variable
- internal validity, 174
- interpreting regression coefficients, *see* linear regression, logistic regression, generalized linear model, *and* multilevel model
- interquartile range (IQR), 70
- intraclass correlation (ICC), 258, 448
- inverse variance, used in `dnorm()` in Bugs, 354–355
- inverse(), 377
- inverse-gamma distribution, why we do not use as prior distribution for variance parameters, 430–434
- inverse-logit function, 80
- invlogit(), 149
- item-response model, 314–321
  - guessing, 319
  - multilevel, 316
  - picture, 315, 317
  - redundant parameters in Bugs, 426
  - two-dimensional, 319
- iterative regression imputation, 539
- iterative simulation, 408–409

## SUBJECT INDEX

615

- Jacobian for nonlinear transformations,  
 409, 430
- Jaycees, 334, 335, 340
- `jitter.binary()`, 89
- jittering, 32, 554
- knowing when to give up, 419
- Lac Qui Parle County, 253
- large regression with correlated errors, 265
- latent-data formulation for logistic  
 regression, 85–86, 120
- latin square, 292, 297, 497–501
- Latinos, hypothetical survey of, 454
- least squares, 39, 387–390  
 augmented data and multilevel  
 model, 396–397  
 weighted, 389
- legislative redistricting, 555
- `length()`, 157
- level, 68
- `lgamma()`, 411
- `library()`, 567
- likelihood, 347, 387–414  
 censored data, 404  
 generalized linear model, 389–390  
 inferential uncertainty and  
 simulation, 392  
 logistic regression, 389  
 picture, 390, 391, 395, 396  
 Poisson regression, 390  
 social networks model, 409–413  
 surface, 390–392
- linear predictor, 79, 305
- linear regression, 31–77, 387–390  
 assumptions, 45–47  
 Bayesian inference, 346  
 binary predictor, 31  
 Bugs model, 360  
 compared to principal component  
 line, 57–58  
 continuous predictor, 32  
 correlation and, 57–59  
 counterfactual interpretation, 34  
 diagnostics, 45–47  
 displaying several, 73–74  
 fitting in R, 38–39  
 general principles for building models,  
 69  
 inferential uncertainty, 40  
 interactions, 34–36  
 interpreting coefficients, 33–34  
 interactions, 35–36  
 least squares, 39  
 matrix notation, 388  
 missing-data imputation, 533–538  
 multiple predictors, 32–34  
 notation, 37–38  
 one predictor, 31–32  
 picture of matrix, 37  
 prediction, 47–49, 70–73, 272  
 predictive interpretation, 34  
 sample size and power calculations,  
 446, 451  
 simulation, 140–148  
 standard error, 40  
 statistical inference, 37–42  
 transformation, 53–77  
 validation, 47–49
- linear transformation, 14, 53–54, 88, 95,  
 122, 294  
 centering a batch of multilevel  
 coefficients, 464–466  
 prior distribution and, 355  
 speeding convergence of Gibbs  
 sampler, 419–427
- linearity, 46
- link function, 109
- `list()`, 143, 350
- `lm()`, 38–39, 402, 565
- `lmer()`, 259–262, 266, 267, 277, 566, 573  
 compared to Bugs, 386  
 limitations, 262, 304, 345  
 logistic regression, 302  
 non-nested model, 289  
 six quick examples, 568–569  
 varying intercepts and slopes, 279–289
- local average treatment effect (LATE),  
 219–220, 229, 233
- `log()`, 59
- log-log model, 64
- `log10()`, 61
- logarithmic transformation, 59–65, 98, 252  
 even when not necessary, 65  
 interpreting regression coefficients, 60  
 interpreting variance parameters, 327  
 picture, 60  
 why we usually don't use base 10,  
 60–61
- logistic distribution, 85  
 Bugs, 384  
 close to normal with standard  
 deviation 1.6, 86, 118, 131
- logistic regression, 79–108  
 binned residual plot, 105  
 Bugs, 381–382  
 latent-data formulation, 384  
 choice models in one and multiple  
 dimensions, 128–131  
 compared to probit, 118, 129  
 computing using `lmer()`, 302  
 deviance, 100, 105

- divide by 4 rule for interpreting coefficients, 82
- graph of coefficient estimates, 306
- graphs of data and fitted curves, 307
- ideal-point model, 314–320
- identifiability, 86, 104, 107
- inference, 83
- interactions, 92–96
- interpreting coefficients, 81–84, 89
- item-response (Rasch) model, 314–320
- latent-data formulation, 85–86, 384
- logit and logit<sup>-1</sup> functions, 80
- missing-data imputation, 533–538
- multilevel, 301–323
  - Bugs, 381–382
  - formula, 302, 303
  - graphing, 304–310
  - interpreting coefficients, 304
  - non-nested, 302–304, 320–321
  - overdispersion, 320–321
- odds ratios, 82–83
- pictures, 80
- plotting data and fitted curve, 80, 84
- predictive comparisons, 81
- prediction, 272
- predictive comparisons, 101–104, 466–473
- propensity score, 207–208
- separation, 104, 107
- simulation, 148
- standard error, 83
- two predictors, 90–92
- varying-intercept, varying-slope model, 310–314
- wells in Bangladesh, 86–92
- logistic-binomial model, 109, 116–117
- overdispersion, 116
- logit, *see* logistic regression
- logit(), 381
- lognormal distribution, 15, 383
- looping indexes in Bugs, 353, 366, 367
- looping, for power calculation, 452
- lowess, 298
- lurking variable, 169
  
- magnetic fields and brain functioning, 481–484
- Mahalanobis distance, 207
- many predictors, multilevel models for, 293–296
- maps, 556–557
- Markov chain Monte Carlo (MCMC), 408–409
- MASS package, 122, 567, 573
- matching, 206–212
  - missing-data imputation, 538
  - propensity score, 207–210, 232
  - R packages for, 230
- matching(), 208
- maternal IQ, 32
- matrix notation, 37–38, 284–287
- matrix of predictors
  - group-level, 252
  - individual-level, 251
- matrix of simulations, 146, 149, 353, 358
- Matrix package, 259, 566
- max(), 381
- maximum likelihood, 388–390
  - censored data, 404–405
  - generalized linear model, 389–390
  - logistic regression, 389
- MCMCpack, 567
- mean(), 56, 359, 382, 477
- mediator, *see* causal inference, controlling for post-treatment variable
- men and women, 337
- mesquite bushes, 68–73
- meta-analysis, 386, 438
- Metropolis algorithm, 385, 408–409
  - picture, 408
  - social networks model, 410
- midterm and final exams, 157
- millet crop, 292
- millimeters, inches, and miles, 53
- min(), 381
- missing at random (MAR), 530, 542
  - impossibility of verifying, 531
- missing completely at random (MCAR), 530
- missing data in R and Bugs, 529
- missing not at random (MNAR), 530
- missing values, not allowed in unmodeled data, 416, 529
- missing-data imputation, 333, 529–543
  - available-case analysis, 532
  - Bugs, 367
  - complete-case analysis, 531
  - congressional elections, 145
  - deterministic or random, 534, 537
  - iterative, 539
  - many variables, 539–540
  - matching, 538
  - model-based, 540–541
  - models, 530–531
  - multilevel data structures, 541
  - one variable, 533–538
  - simple methods, 531–533
  - topcoding, 534
- Mississippi, as poor state, 313
- mixed discrete/continuous data, 126, 537
- mixed effects, *see* random effects



## SUBJECT INDEX

617

- MLWin, multilevel modeling in, 573  
 mnp package, 110  
 model checking, *see* posterior predictive checks *and* residuals  
   using simulation, 158–165, 513–524  
 model comparison, 524–526  
 model extrapolation, 169, 185, 201, 209, 213  
 model-based imputation, 540–541  
 modeled data and parameters, 367  
 modeling the coefficients of a large regression model, 264  
 monotonicity, *see* causal inference, instrumental variables, assumptions  
 month of arrest, 21, 331  
 more than two varying coefficients, 285  
 mothers and children, 31–51, 55–57  
 motivations for multilevel modeling, 6–8, 246–247  
 mtext(), 520  
 multilevel model, 1, 237–342, 463  
   alternative to selecting regression predictors, 294  
   analysis of variance (ANOVA), 490–502  
   assumptions, 247  
   Bayesian inference, 393  
   Bayesian perspective, 346  
   building, 293–296  
     Gibbs sampler, 402  
   building from classical regression, 270  
   causal inference, 503–512  
   combining regression inputs, 293–296  
   compared to classical regression, 463  
   comparison to simpler methods, 310  
   complexity, 246  
   computing, 345–434  
   equivalent sample size, 258, 268  
   factor analysis and, 296  
   fake-data simulation, 363–365  
   fitting in Bugs, 345–386  
   fitting in R, Stata, SAS, and other software, 568  
   fitting using lmer(), 259–262  
   five ways to write, 262–265  
   generalized linear model, 325–342  
   Gibbs sampler, 398–402  
     programming in R, 399–402  
   graphing, 304–310  
   group-level predictors, 265–269, 568  
   group-level variance, superpopulation and finite-population, 459–462  
   how many groups needed, 275–276  
   how many observations per group needed, 275–276  
   imputation at different levels, 541  
   inference for groups with no data, 306  
   instead of comparing significance levels, 482  
   instrumental variables, 509–511  
   interpreting coefficients, 268  
   least squares with augmented data, 396–397  
   logistic regression, 301–323, 568  
     Bugs, 425  
     computing using lmer(), 302  
     non-nested, 302–304, 320–321  
     overdispersion, 320–321  
   matrix notation, 284  
   negative binomial, 332–342  
   non-nested, 289–293, 569  
     Bugs, 380–382, 424  
     identifiability of constant term, 381  
     negative binomial, 332–342  
     redundant parameters, 421–423  
   notation, 251–252  
   ordered logistic regression, 331–332  
   plot of group-level estimates and fitted regression line, 266, 307  
   Poisson (overdispersed), 332–342, 382, 568  
   pooling, 252–259  
   prediction, 272–275, 361–363  
   predictive comparisons, 470  
   prior distribution for variance parameters, 427–434, 499–501  
    $R^2$ , 473–477  
   redundant parameters, 420  
   sample size and power calculations, 447–454  
   six quick examples, 568–569  
   small number of groups, 431–432, 461  
   statistical significance, 271  
   summarizing and displaying, 261  
   understanding and summarizing, 457–486  
   variance parameters, 480–481  
   varying intercepts and slopes, 279–289, 568  
   varying slopes without varying intercepts, 283–284  
   varying the number of groups, 330  
 multilevel modeling  
   costs and benefits, 9, 246–247  
   motivations, 6–8, 246  
   when most effective, 270  
 multilevel structures, 237–249  
   data matrix, 238–240, 243  
   imputation, 541  
 multinomial logit and probit models, 110, 119–124

- storable votes, 120–124
- multiple chains, necessary to monitor
  - mixing, 356–358
- multiple comparisons, why we do not worry about, 22, 484–485
- multiple imputation, combining inferences, 542
- multiplicative model, 59
- multivariate imputation, 539–540
- multivariate normal distribution, 15
- `mvrnorm()`, 143
  
- `n.chains`, number of chains when running Bugs, 356–358, 369
- `n.eff`, effective sample size of Bugs fit, 352, 358
- `n.iter`, number of iterations when running Bugs, 356–358, 369
- `n.sims`, 143
- `n.thin` option in Bugs, 518
- NA, missing value in R and Bugs, 50, 362, 529
- naming inputs, 62
- National Election Study, 73, 311, 342, 385
- National Longitudinal Survey of Youth, 210
- National Supported Work, 231
- natural log, 60–61
- `ncol()`, 361, 519
- negative binomial distribution, 115, 336
  - multilevel model, 332–342
- neighborhoods and crime, 325, 342
- nested subscripts in Bugs, 372
- networks, 297, 333
- New York City schools, 458–459
- Newcomb's speed of light data, 159
- Nicoles, 333, 335, 340
- `nmatch` package, 230
- no data, multilevel inference for groups with, 306
- no pooling, 247, 252–259, 270, 349
  - Bugs model, 360
  - overestimates between-group variation, 253
  - picture, 253
  - problems, 256
  - special case of multilevel modeling, 258
- non-nested models, 241–244, 248–249, 289–293
  - Bugs, 380–381
  - varying intercepts and slopes, 291
  - where to put the intercept or constant term, 381
- nonidentifiability, *see* identifiability
- noninformative prior distribution, 347, 355
  
- Bugs, 354, 355
- nonlinear prediction, 147
- nonparametric regression, 297
- normal distribution, 13–15, 263
  - computing in R, 405
  - estimated regression coefficients, 15, 40, 83
  - inverse-variance parameterization in Bugs, 354–355
  - regression errors, 46
  - truncated, 407
- notation, 263
  - capital letters for matrices, lowercase for vectors and scalars, 167, 252, 376, 383
  - cluster sampling, 447
  - linear regression, 37–38
  - multilevel model, 251–252
  - parameters and probability distributions, 13
  - varying intercepts and slopes, 284
- `nrow()`, 519
- number of iterations when running Bugs, 356–358, 369
- number of observations and groups needed, 275–276, 278
- number of sequences when running Bugs, 356–358, 369
- numerical optimization in R, 405
- nutrients and cancer, 294
- NYPD stops, *see* police stops
  
- O'Connor, Sandra Day, 318
- observational study, *see* causal inference
- odds ratios and logistic regression, 82–83
- offset, in Poisson regression, 112, 326, 382
- Ohio, as intermediate state, 313
- Olympics, 248, 485
- omitted variable, 169
- omniscience, 195
- one-way analysis of variance, 494
- open-ended modeling in Bugs, 370–372
- OpenBugs, 11, 565, 574, *see also* Bugs
- `optim()`, 405, 413
- optimal design, 455
- `options()`, 561
- `order()`, 519
- ordered and unordered categorical outcomes, 119, 123
- ordered logistic model, 119–124
  - fitting in Bugs, 383
  - multilevel, 331–332
  - storable votes, 120–124
- outcome, 37, 251
- `outer()`, 411
- overdispersion, 21, 114–116, 320

## SUBJECT INDEX

619

- adjusting standard errors, 115, 117
- groups in the social network, 338
- multilevel Poisson model, 335–336, 382, 409–413
- simulation, 150
- variance components and, 325
- overlap, *see* causal inference
- pain scores, observed vs. expected, 558
- panel-corrected standard errors, 248
- par(), 305
- parameter expansion, *see* redundant parameters
- parameters saved from Bugs, 350, 356
- partial pooling, 252–259, 394, 477–480
  - Bayesian, 394
  - formula, 253, 258, 269, 477
  - graph, 479
  - group-level predictors and, 269
  - picture, 253
  - plotting data and fitted lines, 257, 266
  - set of regression predictors, 295
  - summarizing a fitted multilevel model, 477–480
- partisan bias, 555
- paste(), 353
- pch, 43
- $p_D$ , effective number of parameters in a Bayesian inference, 525
- phase diagram for decision analysis, 130
- plot(), 43, 350
- plots of replicated datasets, 160, 163
- pmin(), 478
- pnorm(), 404
- points(), 43
- Poisson model, 16, 109–116, 335
  - checking using simulation, 161–163
  - compared to binomial, 112
  - exposure, 111–113
  - interpreting coefficients, 111
  - multilevel, 325–331
  - offset, 326
  - overdispersion, 114–116
  - police stops, 112–116
  - zero-inflated, 126–127
- police stops, 5–6, 21, 112–116, 325–331, 342
  - Bugs model, 382
  - graph, 328
- political ideology, 73
- political party identification, 73–74
- pollution, 76
- polr(), 110, 122
- pooling, *see* complete pooling, no pooling, partial pooling
- pooling factor, 478–480
- posterior distribution
  - picture, 395, 396
  - programming in R, 411
  - social networks model, 409
- posterior predictive checks, 158–165, 513–515
  - data display for dog example, 516, 523
  - numerical summary, 521
  - time plot, 519, 520, 522
  - using Bugs, 518
  - using R, 518
- posterior uncertainty, 149
- postprocessing Bugs output, 359
- poststratification, 178, 181, 206, 301–310
  - formula, 301, 308
  - R code, 308
- potential outcome, 168, 171, 183, 186, 189, 191, 219
  - close substitutes, 171–172
  - interpreting regression coefficients, 34
- pow(), 355
- power calculation, 437–455
  - 2.8 standard errors, 441
  - classical, 439–447
  - general concerns, 439
  - inference for continuous outcomes, 443–447
  - inference for linear regression, 451
  - inference for proportions, 439–443
  - inference for regression coefficients, 446–447
  - inherently speculative, 445, 447
  - multilevel models, 447–454
  - pictures, 440, 441
  - unequal sample sizes, 443
- pre-election polls, 560
- precinct, 325
- predict(), 48, 115, 208, 535
- prediction, 47–49, 68–73
  - Bugs, 361–363
  - interpreting regression coefficients, 34
  - model checking and, 513–515
  - multilevel model, 272–275, 361–363
  - new observations and new groups, 272–275, 361–363
  - nonlinear, 274
- predictive checks, *see* posterior predictive checks
- predictive comparison, 81, 101–104, 167, 168, 466–473, 485
  - comparing models, 472–473
  - formula, 466
  - general approach, 468
  - general notation, 103
  - graph, 467, 468

- interactions and, 103
- model summary, 471
- predictive simulation, 140, 147–151
  - binomial distribution, 149
  - generalized linear models, 148–151
  - latent logistic distribution, 149
  - linear regression, 140–148, 152
  - model checking and, 158–165, 513–524
- predictive standard deviation, 274
- predictive uncertainty, simulation of, 140
- predictors, as distinguished from input variables, 37, 466
- presidential elections, 3–4, 79–84, 294, 301–314, 493, 557, 560
- principal component line, 57
- principles of modeling in Bugs, 366–369
- print(), 350
- prior distribution, 143, 345–348, 413, 427–434
  - Bugs, 354
  - effect on posterior distribution, picture, 430, 432
  - informative, 392–393
    - scale, 430
  - inverse-gamma, why we do not use, 430–434
  - noninformative, 347, 354, 355
  - picture, 395, 396
  - provisional nature of, 347
  - scale, 355
  - uniform, 428–429
  - variance parameters, 432–434, 499–501
  - weakly informative, 431–432
  - Wishart model, 377
- prison sentences, example for predictive comparisons, 470
- probability, *see* distribution
- probability models, simulation of, 137–140
- probability of a tied election, 148
- probit model, 109, 118–119
  - compared to logit, 118, 129
- programming in R, 567
- Progresá, 508–509
- propagating uncertainty, 142, 152
- propensity score matching, *see* causal inference
- provisional nature of prior distributions, 347
- psychological experiment of pilots on flight simulators, 289–290
- quantile(), 141, 359
- quasibinomial family, 117
- quasipoisson family, 115
- quick tips, 547–549
- R, 10–11, 298, 565, 573
  - abline(), 353, 520
  - apply(), 44, 353, 477
  - array(), 308
  - as.bugs(), 413
  - as.bugs.array(), 400
  - as.vector(), 348
  - attach.bugs(), 305, 352, 358
  - Bayesian inference for social networks model, 409–413
  - brlr package, 104
  - bugs(), 350, 567
  - c(), 350
  - calling Bugs from, 350–352
  - cbind(), 43, 146, 157, 361, 529
  - censored data, 404–408
  - coef(), 43, 156, 260, 267, 280, 352
  - colMeans(), 520
  - colors(), 43
  - colSums(), 411
  - console, 565
  - curve(), 43, 353
  - data.frame(), 48, 140, 452, 535
  - default values in functions, 452
  - digits, 561
  - dim(), 147
  - dimnames(), 400
  - display(), 38–39, 565
  - dnorm(), 404
  - expected(), 123
  - expression(), 148, 353
  - factor(), 255, 349
  - fitted(), 158
  - fixef(), 260, 280
  - function(), 350, 401
  - GAMM package, 567
  - Gibbs sampler, 399–402
    - censored data, 406–408
  - glm(), 79, 110, 565
  - global variables, 400, 412
  - graphics, 562
  - graphics window, 565
  - graphing models fit in Bugs, 352
  - help, 405, 565, 567
  - hett package for robust regression, 110, 124, 133, 567
  - hist(), 137, 536, 562
  - I(), 215, 538
  - ifelse(), 126, 150, 384, 403, 534
  - impute(), 535
  - install.packages(), 567
  - invlogit(), 149
  - jitter.binary(), 89
  - length(), 157

## SUBJECT INDEX

621

- lgamma(), 411
- library(), 411, 567
- list(), 143, 350
- lm(), 38–39, 402, 565
- lmer(), 259–262, 266, 267, 277, 566, 573
  - limitations, 262, 304, 345
  - logistic regression, 302
  - non-nested model, 289
  - six quick examples, 568–569
  - varying intercepts and slopes, 279–289
- log(), 59
- log10(), 61
- log=TRUE option, 405
- MASS package, 122, 567, 573
- matching(), 208
- Matrix package, 259, 566
- MCMCpack package, 567
- mean(), 56, 359, 477
- mnp package, 110
- mtext(), 520
- mvrnorm(), 143
- NA, 50
- ncol(), 361, 519
- nmatch package, 230
- nrow(), 519
- optim(), 405, 413
- options(), 561
- order(), 519
- outer(), 411
- par(), 305
- paste(), 353
- pch, 43
- plot(), 43, 350
- pmin(), 478
- pnorm(), 404
- points(), 43
- polr(), 110, 122
- predict(), 48, 115, 208, 535
- print(), 350
- probit family, 118
- programming, 567
- quantile(), 141, 359
- quasibinomial family, 117
- quasipoisson family, 115
- R2WinBUGS, 565
- ranef(), 260, 280
- range(), 352, 520
- rbinom(), 137, 149
- read.dta(), 411
- read.table(), 49, 348
- rep(), 452
- replicate(), 139, 147
- replicated data for predictive checking, 518
- return(), 401
- rnegbin(), 150
- rnorm(), 106, 141, 155, 356, 401, 407
- rnorm.trunc(), 407
- rowMeans(), 180
- rowSums(), 147, 411
- rpois(), 150
- runif(), 150, 353, 356
- rwish(), 377
- sample(), 138, 278, 418, 452, 534
- sapply(), 151
- save(), 362
- sd(), 56, 462
- se.coef(), 156, 565
- se.fixef(), 261
- se.ranef(), 261
- sem package, 223
- sigma.hat(), 273, 565
- sim(), 43, 142, 143, 392, 565
- sorting, 519
- subset option in lm() and glm(), 107, 126, 538
- sum(), 147
- summary(), why we don't use, 39
- table(), 353
- tlm(), 124, 133, 567
- tsls(), 223
- unique(), 348
- updating functions for Gibbs sampler, 399–402
- var(), 477
- R functions, 44, 139, 147, 151, 404, 534, 535
- $R^2$ , 41, 49, 62, 485
  - adjusted, 475
  - Bayesian definition, 475
  - classical definition, 474
  - computation, 476
  - each level of a model, 474
  - interpretation for model with no constant term, 349
  - multilevel, 473–477
  - pictures, 42
  - why we do not define in terms of model comparison, 474
- R2WinBUGS package, 565
- $\hat{R}$  for summarizing convergence of Bugs fit, 352, 358, 369
- radon, 3, 36, 252–283, 348–369, 480–481
- random effects, 244–246, 259
  - many definitions, 245, 248
  - R, 260
  - superpopulation inferences and, 461
  - why we avoid the term, 2, 245
- random imputation, 534, 537

- randomized experiment, 171–181, 183, *see also* causal inference  
 ranef(), 260, 280  
 range(), 352, 520  
 Rasch model, 315–320  
 ratings, 298  
 ratio of parameters, 152  
 raw (unscaled) parameters, 377  
     compared to adjusted parameters, 423  
 rbinom(), 137, 149  
 read.dta(), 411  
 read.table(), 49, 348  
 recall, in social networks survey, 339  
 red states and blue states, 310–314  
 redundant parameters, 419–427  
     additive, 316, 326, 336, 382, 412, 419–423, 464–466  
     future implementations, 427  
     item-response and ideal-point models, 316  
     multiplicative, 424–427  
         Bugs, 424, 425  
         social network model, 336  
 reference condition, in classical no-pooling regression, 349  
 reference model, 347  
 regression, *see* linear regression, logistic regression, generalized linear models  
 regression coefficients, graph, 337, 341  
 regression discontinuity, 212–215  
 regression to the mean, 57–59  
 rejection, not the goal of model checking, 524  
 rep(), 452  
 repeated measurements, 241–243  
     graph, 450  
 replicate(), 139, 147  
 replicated data, for predictive checking, 514  
 replicated datasets, plotted, 160, 163  
 residuals, 40, 97–101  
     binned, 97–101, 559  
     distinguished from errors, 387  
     plot, 47, 48, 97, 114, 558  
     plot vs. predicted values, not vs. observed values, 157, 158  
     social network model, 341–342  
     square root, for Poisson model, 341–342  
     standard deviation of, 41  
 return(), 401  
 rnegbin(), 150  
 rnorm(), 106, 141, 155, 356, 401, 407  
 rnorm.trunc(), 407  
 robit regression, 124–125, 133, 320  
     Bugs, 384  
     generalization of logit and probit, 125  
     latent-data formulation, 384  
     picture, 124  
 robust regression, 110, 131  
 rodents, 106, 248, 322  
 rowMeans(), 180  
 rowSums(), 147, 411  
 rpois(), 150  
 runif(), 150, 353, 356  
 rwish(), 377  
  
 S and S-Plus, *see* R  
 sample size and interactions, 438  
 sample size calculation, 437–455  
     2.8 standard errors, 441  
     classical, 439–447  
     general concerns, 439  
     inference for continuous outcomes, 443–447  
     inference for linear regression, 451  
     inference for proportions, 439–443  
     inference for regression coefficients, 446–447  
     inherently speculative, 445, 447  
     multilevel models, 447–454  
     pictures, 440, 441  
     unequal sample sizes, 443  
 sample(), 138, 278, 418, 452, 534  
 sampling, design for, 437–455  
 sapply(), 151  
 SAS  
     code for matching, 230  
     multilevel modeling in, 570–571  
 save(), 362  
 scale of prior distribution, 355  
 scale-up model for estimation in a social network, 333  
 scaled inverse-Wishart distribution, 284–287, 298, 376–380  
 Scalia, Antonin, 318  
 scaling of predictors, 53  
 scatterplot  
     advice, 553–559  
     data and regression lines superimposed, 35  
 sd(), 56, 460, 462  
 se.coef(), 156, 565  
 se.fixef(), 261  
 secret weapon, 73–74  
     income and voting, 311  
     pictures, 19, 74, 84  
 selection bias, 168, 231, *see also* causal inference  
 selection on observables, *see* ignorability  
 sem package, 223

## SUBJECT INDEX

623

- separation in logistic regression, 104, 107
- Sesame Street, 196, 216–220, 231, 509–511
- sex ratio of births, 27, 137–139
- shrinkage and partial pooling, 477
- `sigma.hat()`, 273, 565
- significance, *see* statistical significance
- significant digits and uncertainty, 561
- `sim()`, 43, 142, 143, 392, 565
- simple and complex models, 416
- simulation, 19–20, 137–166
  - combined with analytic calculations, 148
  - comparing simulated to actual data, 158–165
  - compound models, 150–151, 537–538
  - coverage of confidence intervals, 156
  - displaying uncertainty in a fitted model, 149
  - fake data, 50, 155–158
  - generalized linear models, 148–151
  - how many draws are needed, 153
  - logistic regression, 148
  - matrix of simulated parameters and predictions, 146, 149
  - nonlinear predictions, 144–148
  - overdispersed models, 150
  - posterior predictive checks, 158–165, 513–524
  - predictive, 148–151
  - probability models, 137–140
  - regression inferences, 140–148
  - replicated datasets, plotted, 160, 163
  - saved as vectors and matrices, 353, 358
  - why necessary, 141
- slope, *see* varying intercepts and slopes
- small multiples plot, 255, 257, 266, 291, 560
  - logistic regression, 307
- small-area estimation, 301–310
- smoking, 36, 241–243
  - data matrix, 242, 243
- Social Indicators Survey, 529–543
- social networks, 332–342
  - Bayesian inference, 409–413
  - graph of data, 335
  - group sizes, 339
  - predicted from demographics, 337
  - residuals, 341–342
- soft constraint, 257
- software, 565–574, *see also* R, Bugs, Stata, SAS, SPSS, AD Model Builder, HLM, MLWin
  - data and code for examples, 11
  - getting started, 565
  - multilevel modeling, 573
  - S and S-Plus, *see* R
  - WinBugs and OpenBugs, *see* Bugs
- speed dating experiment, 322, 323
- speed of light, 159
- splines, 298
- split-plot latin square, 498–501, 509
- SPSS, multilevel modeling in, 571
- square root transformation, 249, 535
- standard deviation, *see* variance
- standard error, 17, 40
  - picture, 40
  - proportions, 17
- standardizing predictors, 54–57, 96
- Stata
  - multilevel modeling in, 569–570
  - reading in data from, 50, 411
- state-level opinions from national polls, 4–5, 301–310, 493–494
  - Bugs model, 381
- statistical inference, 16–17, 37–42
  - graph of uncertainty, 40, 83
  - measurement error model, 16
  - sampling model, 16
  - standard error, 40, 83
- statistical significance, 42, 69, 83, 94
  - limitations of, 481–484
  - multilevel model, 271
  - problems with, 22–23
  - sample size and power, 440
- stochastic learning in dogs, 515–524
  - Bugs model, 517, 521–524
  - model comparison, 526
- stop and frisk, *see* police stops
- storable votes, 120–124, 331–332, 386
  - Bugs model, 383
  - data and fitted curves, 121
- strategy of debugging, 416
- structural equation modeling, 231
- subclassification, 206–207, 229
- subset option in `lm()` and `glm()`, 107, 126, 538
- subsetting data, 326, 357, 547
  - speeding computation, 418
- sum of squares
  - analysis of variance, 488
  - least squares estimation, 387
- `sum()`, 147
- superpopulation, 167, 173
  - analysis of variance, 491, 500
  - variance, 459–462
    - computing in Bugs, 460
    - graph, 461
- Supreme Court voting, ideal-point model
  - for, 317
- survey design, 437
- survey weighting, 301–310

- switches, 165
- t distribution, 124, 131, 372, 428
- table(), 353
- tables, 563
- teachers, effect of, 459
- teaching evaluations and beauty, 51
- test summary, 513–515
  - graphical, 160, 163, 516, 519, 522, 523
  - numerical, 23, 159, 161, 521
- text editor, 565
- thinning in Bugs, 418, 518
- thresholds, for ordered logit or probit, 119–120
- tied election, probability of, 148
- time series, 297
  - checking a fitted model, 163–165
  - cross-sections, 243–244, 248
- tlm(), 124, 133, 567
- tobit model, 126, 132
- topcoding, for missing-data imputation, 534
- traffic accidents, 110–111
- transformation, 53–77, 548
  - idiosyncratic, 65
  - linear, 53–54
  - logarithmic, 59–65, 98, 252
    - interpreting variance parameters, 327
  - square root, 65, 249, 535
- treatment effect, *see* causal inference
- true values in fake-data simulation, 155, 363
- truncated normal distribution, 407
- tsls(), 223
- twins, 138
- two-factor experiment, 289–290
- two-level classical regression, 240, 248, 270
- two-stage least squares, *see* causal inference, instrumental variables
- two-stage model for mixed discrete/continuous data, 126, 537
- two-way analysis of variance, 495, 496
- U.S. Census, 277, 301, 308
- UCLA, 332
- Ulysses, 339
- Umacs, 337, 410–413, 567
- uncertainty, as distinguished from variability, 457–459
- uncontested elections, as missing data, 145
- underdispersion, 21, 22
- unemployment series, graph of data and replications, 163–165
- unexpected patterns, discovering through graphs, 551
- unexplained variance, *see*  $R^2$
- unique(), 348
- units, 37, 251, 553
- unmodeled data and parameters, 364, 367, 378
- unordered categorical regression, 124
- updating functions in R, 399–402
- uranium, as county-level predictor in radon model, 266
- utility theory, 128
- validation, 47–49
- validity, 45, 174
- value added by schools, 458–459, 485
- value function, 128
- value of a statistical life, 197
- var(), 477
- variability, as distinguished from uncertainty, 457–459
- variance
  - explained and unexplained, 41, 473–477
  - group level, 270
  - models for, 297
  - multiple error terms, 264
  - non-nested models, 290, 291
  - predictive, 274
  - ratio of between to within, 258
  - residual, 41
  - superpopulation and
    - finite-population, 459–462
- varying intercepts, *see* multilevel model
- varying intercepts and slopes, 1, 237, 279–289, 549
  - Bayesian perspective, 346
  - Bugs, 375–379
  - computing using lmer(), 282
  - graph, 450
  - group-level predictors, 280, 379–380
  - interactions, 282–283
  - logistic regression, 310–314
  - non-nested model, 291
  - notation, 284
  - pictures, 238
- varying slopes without varying intercepts, 283–284
- vector of simulations, 353, 358
- vector-matrix notation, 37–38, 284–287
  - Bugs, 361
- Vietnam War draft lottery, 225–226, 230
- visual and numerical comparisons of replicated to actual data, 164
- voting and income, 79–84, 105, 107, 310–314



## SUBJECT INDEX

625

- Wald estimate for instrumental variables,  
219, 221
- wavelets, 298
- weight
- age and, 75
  - example of a lognormal distribution,  
15
  - height and, 74, 402–408
- weighted average, 19
- weighted least squares, 389
- wells in Bangladesh, 86–92, 105, 193
- arsenic levels, 90
  - choice models, 127–131
  - map, 87
- when does multilevel modeling make a  
difference, 247
- WinBugs, *see* Bugs
- WinEdt, 565
- Wishart distribution, 284–287, 298,  
376–380
- world wide web connect times, 492–493
- $\chi^2$ test, 25–26, 114
- z-score, 50, 54
- zero-inflated Poisson model, 126–127
- zinc and HIV, 443–447