Index

D-optimal design, 247
G-optimal design, 246
χ-squared distance, 186
σ-algebra, 21

, 8, 20

a posteriori, 28
a priori, 28
Abel sum, 470, 478
absolutely continuous, 40, 61, 181
action gap, 59
action space, 448
adapted, 28
adversarial bandit, 12, 143, 145
affine hull, 425
anytime, 91
Assouad's method, 195
asymptotic optimality, 125, 134, 200, 363, 375, 415, 445

Bachelier-Levy formula, 127
bandits with expert advice, 216
Bayes law, 29
Bayesian regret, 58, 380, 381
Bayesian upper confidence bound, 415
belief space, 469
Bellman operator, 455
Bellman optimality equation, 453, 454
Bernoulli bandit, 56, 117
Bernoulli distribution, 34
Bernstein's inequality, 77, 82, 473
best arm identification
  fixed confidence, 368
beta distribution, 387
bias-variance tradeoff, 213
Borel, 22
Borel space, 46
Bregman divergence, 303
Brownian motion, 126, 351

canonical bandit model, 60, 190
cascade model, 353
categorical distribution, 463
Catoni's estimator, 112
cell decomposition, 425
Cesàro sum, 453, 470, 478
Chernoff bound, 130, 362
click model, 353
completement, 21
concave, 286
conditional entropy, 412
conditional expectation, 33
conditional probability, 28
conjugate prior, 386
consistent policy, 198, 276
context, 14
controlled Markov environment, 475
convex optimization, 302
counting measure, 40, 188, 190
covering, 237, 244
Cramer–Chernoff method, 70, 164, 239, 240, 243
cumulant generating function, 138
cumulative distribution function, 33
cumulative generating function, 72

data-dependent bound, 220, 222
degenerate action, 425
descriptive theory, 63
design matrix, 246
deviation matrix, 477

diameter, 451
differential value function, 453, 454
discount factor, 348
discounting, 348, 395, 400
domain, 285
dominated action, 425
dominating measure, 40
doubling trick, 91, 156, 266, 313
dual norm, 308

easy partial monitoring problem, 424
empirical Bernstein, 107, 111
empirical process, 84
empirical risk minimization, 222
entropy, 179, 187, 412
environment, 10
epoch-greedy algorithm, 225
essentially smooth, 288
essentially strictly convex, 288
Exp3, 149, 214, 295, 297, 302, 316, 331, 343, 345, 423
Exp3-IX, 162, 205, 208, 331, 350
Exp3.P, 350
Exp3.S, 350
Exp4, 218, 344
expectation, 31
expectation operator, 41
exponential family, 117, 137, 139, 202, 370, 387, 415
exponential weighting, 149
exponential weights, 335
Exponential weights algorithm, 154
extended real line, 285

feasible, 455
feedback matrix, 422
Fenchel dual, 286, 335
filtered probability space, 28
filtration, 28, 83, 138, 163, 239, 245, 260, 390
finite additivity, 21
first order bound, 167
first-order optimality condition, 290
Fisher information, 193
Fixed Share, 350
follow the leader, 303, 314
follow the perturbed leader, 159, 221, 335, 416
follow the regularized leader, 302
Fubini's theorem, 41, 65
full information, 154, 294, 350
functional, 41
fundamental matrix, 477

gain, 452
generalized linear model, 233
Gittins index, 350, 394
globally observable, 427
gradient descent, 304
Gram matrix, 276

Hahn decomposition, 32
hard partial monitoring problem, 424
Hardy–Littlewood, 397
heavy tailed, 77
Hedge, 154
Hellinger distance, 186
high probability bounds, 161
history, 10
Hoeffding's inequality, 83, 362
Hoeffding's lemma, 77, 131, 154, 245
Hoeffding–Azuma, 244, 437
hopeless partial monitoring problem, 424

image, 441
immediate regret, 59
implicitly normalized forecaster, 318
importance-weighted estimator, 147, 148, 166, 222, 310
independent events, 29
index, 395
index policy, 394
indicator function, 24
information directed sampling, 417
instance-dependent bound, 222
integrable, 31
Ionescu Tulcea, 62
isomorphic measurable spaces, 46

Jensen's inequality, 39, 287
John's theorem, 249

joint distribution, 23
Jordan curve theorem, 442
Jordan-Brouwer separation theorem, 436, 442

Kahan's algorithm, 149
kernel, 441
kernel trick, 232
Kiefer–Wolfowitz Theorem, 247
Kullback-Leiblier divergence, 179

Laplace's method, 240
Laurent series, 478
law, 22
law of the iterated logarithm, 108, 121, 260
lazy mirror descent, 313
Le Cam's method, 194
learner, 10
learning rate, 149, 163, 220, 295, 296, 303, 304, 345, 433
  adaptive, 313, 316, 317
least-squares, 236
Lebesgue σ-algebra, 38
Lebesgue integral, 31
Legendre function, 288, 303, 335
light tailed, 77
linear subspace, 441
link function, 233
locally observable, 427
log partition function, 139, 387
log-concave, 297
loss matrix, 422

Markov chain, 433, 450
Markov kernel, 48
Markov policy, 450
Markov process, 51
Markov property, 469
martingale, 49, 78, 82, 239
martingale difference process, 239
martingale noise, 239
maximal end-component, 483
maximal inequality, 121
maximum end-component, 484
measurable set, 21
measurable space, 21
measure, 22
measure theory, 37
median-of-means, 111
memoryless deterministic policy, 450
memoryless policy, 450
metric entropy, 244
minimax optimal, 174
minipage, 120
mirror descent, 154, 302, 335, 345
MOSS, 117, 120
multiclass classification with bandit feedback, 223
multitask bandit, 271, 332
mutual information, 412

negentropy, 345
neighboring actions, 425
nonoblivious, 155, 313
nonparametric, 56
nonstationary, 155
nonstationary bandit, 64
null set, 38

oblivious, 155, 308, 313
one-armed bandit, 67, 118, 390
online gradient descent, 304
online learning, 16, 222, 262, 302
online linear optimization, 302
online-to-confidence set conversion, 261
operator, 41
optimal design, 246
optimal experimental design, 375
optimal stopping, 391
optimal value function, 454
optimism bias, 107
optimization oracle, 221, 339
optional stopping theorem, 50
orthogonal complement, 441
outcome space, 19, 383

packing, 244
parameter noise, 322
parameter space, 383
parametric, 56
Pareto optimal, 177

Pareto optimal action, 425
partial monitoring, 421
partially observable Markov Decision Processes, 469
peeling device, 121
permutation, 63, 352
Pinsker's inequality, 130, 138, 183, 186, 187, 200, 308
point-locally observable, 443
polar, 285, 321, 324
policy, 10, 60
policy iteration, 471
policy regret, 155
POMDPs, 469
position-based model, 353
potential function, 303
predictable, 28
prediction with expert advice, 154
preimage, 20
premetric, 130
prescriptive theory, 63
prior, 381
prior variance, 386
probability kernel, 48, 383
probability measure, 21
probability space, 22
process, 28
product measure, 61
projective, 47
pseudo regret, 64
pseudo regret, random, 206
pushforward, 22

quadratic variation, 167

Radon-Nikodym derivative, 39, 61, 65, 190
random regret, 64
random variable, 20
ranked bandit model, 362
reactive, 155
reduction, 320, 368
regret, 10
  nonstationary, 344
  tracking, 344
regret decomposition lemma, 59
regularizer, 303
relative entropy, 129, 140, 179, 288
restless bandit, 350
retirement policy, 68, 390
ridge regression, 236
right stochastic, 433
right stochastic matrix, 433, 448

semibandit, 330
separation oracle, 298, 457
Sequential Halving algorithm, 374
signal variance, 386
signed measure, 22
similarity function, 216
simple function, 31
simple regret, 366
Snell envelope, 391
span, 453
standard optimal stopping, 391
state space, 448
static experts, 221
stationary transition matrix, 453, 476
stochastic optimization, 374
stochastic process, 47
stopping time, 50, 95, 196, 260, 271, 368, 372, 377, 390, 451
strictly convex, 286
strongly connected component, 483
structured bandit, 57
sub-σ-algebra, 21
submartingale, 49
suboptimality gap, 59
sufficient statistic, 139, 387
supermartingale, 49, 243, 260
supervised learning, 216
support, 4, 57, 117, 146, 222, 250, 297, 334, 474
support function, 335, 341
supporting hyperplane, 290

Taylor's theorem, 240
Thompson sampling, 64, 117
total variation distance, 186
Track-and-Stop algorithm, 371

transductive learning, 223
transition matrix, 452
trivial events, 44
trivial partial monitoring problem, 424
uniform exploration, 367
union bound, 72
unnormalized negentropy, 293, 305, 315, 333, 345
unnormlized negentropy, 288
unrealizable, 272, 323
unstructured bandit, 57
unstructured bandits, 198
value iteration, 455
vanishing discount approach, 470
Varaiya's algorithm, 401
von Neumann-Morgenstern theorem, 63
weak neighbor, 443
worst case regret, 174