



# データサイエンス系 検定まとめ

生成AI時代のDS検定のあり方は？

ChatGPTに聞いてまとめました

庭師



# 日本で主要なデータサイエンス系検定

| 検定・資格名         | 主催                  | レベル感             | 特徴                                | 主な対象者            |
|----------------|---------------------|------------------|-----------------------------------|------------------|
| 統計検定®          | 日本統計学会・統計質保証推進協会    | 4級～1級、データサイエンス基礎 | 日本で最も有名な統計資格。統計理論から応用まで体系的。       | 学生、研究者、社会人実務者    |
| データサイエンス検定™    | 日本統計学会＋JDLA         | 初級～中級            | 統計＋データ処理＋機械学習＋活用事例。2021年開始の新しい検定。 | 社会人・学生全般、文系も可    |
| G検定(ジェネラリスト検定) | 日本ディープラーニング協会(JDLA) | 初級               | AI/ディープラーニングの知識面。数式不要。業界動向や倫理も出題。 | AIに関心のある社会人、文系含む |
| E資格(エンジニア資格)   | 日本ディープラーニング協会(JDLA) | 上級               | Python＋深層学習の実装スキルを問う。認定プログラム受講必須。 | AIエンジニア、研究者      |
| 応用情報技術者試験      | IPA(情報処理推進機構)       | 中級               | IT全般知識。AI・統計も一部出題。国家資格。           | ITエンジニア、社会人実務者   |

## 入門(文系・学生向け)

統計検定 3級・4級  
データサイエンス検定 基礎  
G検定

## 実務・社会人レベル

統計検定 2級  
データサイエンス検定  
応用情報技術者

## 専門家レベル

統計検定 準1級・1級  
E資格



# Kaggle titanicコンペでスコア0.80が出せる人には どの検定がお勧めですか？

- Kaggle Titanic で LB 0.80 以上を出せるなら、すでに以下の力を持っていると考えられます：
  - pandas / scikit-learn / 可視化ライブラリを使った前処理・EDA能力
  - 欠損値処理・特徴量エンジニアリングの発想力
  - モデル選択(ランダムフォレスト・GBDT・アンサンブルなど)
  - クロスバリデーションを使った評価スキル
- つまり「データ分析の実践力はすでに初級～中級を超えている」状態です。

# おすすめの検定

- **本格派に伸ばすなら**

- **統計検定 2級 ~ 準1級**

- 数学・統計の裏付けを体系的に固めたい場合。
    - 特徴量エンジニアリングやモデル解釈力をさらに伸ばせます。

- **幅広さを示すなら**

- **データサイエンス検定(JDLA/統計検定協会)**

- 統計 + 機械学習 + 活用事例を横断的にカバー。
    - Kaggle経験者ならかなり有利。企業評価も広がしやすい。

- **AI志向なら**

- **G検定**

- 「ディープラーニングを使う/議論する」ための基礎知識。
    - 文系にも取れるけれど、Kaggle経験があれば楽勝レベル。

- **E資格(さらに進みたい人向け)**

- Python + 深層学習実装スキルを問う実務資格。
    - Titanicでの特徴量づくり・CV管理ができている人なら、講座受講後に挑戦する価値あり。

# 統計検定®(統計検定協会)

- 日本統計学会が監修。最も広く知られる統計資格。
- レベル：
  - 1級:大学院レベル。かなり難易度高め(研究者レベル)。
  - 準1級:大学で統計を専攻した人向け。
  - 2級:大学基礎～実務統計。社会人がよく受験。
  - 3級:高校～大学教養レベル。
  - 4級:高校レベル。
  - 統計検定 データサイエンス基礎:高校生や文系社会人でも取りやすい入門。
- 実務での統計力を客観的に示すのに有効。

# 統計検定®(統計検定協会): 1

- ・主催: 統計質保証推進協会(監修: 日本統計学会)
- ・日本で最も広く認知されている統計系の資格試験。
- ・統計の基礎理論からデータサイエンスへの応用まで、段階的に習得できる。
- ・大学・企業でも履歴書評価されやすい。

| 級・区分       | レベル     | 内容                             | 対象者の目安      |
|------------|---------|--------------------------------|-------------|
| 1級         | 大学院レベル  | 数理統計・応用統計・統計応用(経済・工学など)を専門的に問う | 統計学研究者、専門職  |
| 準1級        | 大学上級    | 数理統計＋応用力。統計的推測の高度な理解が必要        | データ分析高度実務者  |
| 2級         | 大学基礎～実務 | 確率・推測統計・回帰分析・分散分析など。実務活用可能     | 社会人、大学生     |
| 3級         | 高校～大学教養 | 確率分布、標本調査、基礎的な推定と検定            | 初学者、文系学生    |
| 4級         | 高校基礎    | 平均・分散、グラフ、データの整理など             | 高校生、数学基礎レベル |
| データサイエンス基礎 | 入門      | データ分析リテラシー、表やグラフの読み取り          | 社会人一般、文系でも可 |

# 統計検定®(統計検定協会):2

## ・ 出題範囲(共通)

- ・ 記述統計学:平均・分散・標準偏差・相関係数
- ・ 確率論:確率分布(二項分布、正規分布など)
- ・ 推測統計:推定・仮説検定
- ・ 回帰分析・分散分析
- ・ 応用分野:品質管理、経済・社会調査、医療統計など

## ・ 試験方式

- ・ CBT方式(コンピュータ受験)。会場の端末で回答。
- ・ マークシート／記述(級による)。
- ・ 試験時間:50分(4級)～120分(1級)。
- ・ 合格基準:正答率 70%前後(級により異なる)。

## ・ メリット

- ・ 統計スキルの「客観的証明」になる。
- ・ 学術・実務の両面で評価されやすい。
- ・ 大学入試・推薦入試での活用実績もあり。
- ・ 企業(特に製造業・金融・IT)での人材育成に利用されている。

## ・ どの級を目指すべきか？

- ・ 統計初心者／文系 → 3級 or データサイエンス基礎
- ・ 大学生・社会人分析担当 → 2級
- ・ 実務経験＋研究志向 → 準1級～1級
- ・ Kaggle Titanicで0.80を出せるレベルなら、統計検定 2級 → 準1級 を目指すと「実務＋理論」のバランスが取れておすすめです。

# データサイエンス検定™ (JDLA × 統計検定協会): 1

- 主催: 一般社団法人データサイエンス教育協議会 (JDLAと統計検定協会が連携)
- 開始: 2021年
- 目的: データサイエンスの基礎知識と活用スキルを持っていることを客観的に証明する。
- 特徴: 統計・プログラミング・機械学習を横断的にカバーする、日本初の「総合型」検定。
- 出題範囲
  - 統計・機械学習・プログラミング・応用事例 と広く浅くカバー。
- ITパスポートのデータ版のような位置づけ。文系でも挑戦しやすい。
- データサイエンス検定™には「級」や「レベル分け」はありません。合格・不合格のみ。
  - 単一レベルの検定 → 「データサイエンスの基礎力があるか」を判定する1種類の試験のみ。
  - 難易度は 大学初級～社会人実務の基礎レベル。

# データサイエンス検定™ (JDLA × 統計検定協会): 2

## ・ 出題範囲

- ・ 大きく4つの領域から出題されます:
- ・ **データサイエンスの基礎**
  - ・ データの種類、収集、前処理、可視化、要約統計量
- ・ **統計学の基礎**
  - ・ 確率、推定、検定、回帰分析、分散分析
- ・ **機械学習の基礎**
  - ・ 教師あり学習(回帰・分類)
  - ・ 教師なし学習(クラスタリング、主成分分析)
  - ・ モデル評価(過学習、クロスバリデーション)
- ・ **データサイエンスの活用事例**
  - ・ 製造業、金融、医療、マーケティングなどの応用
  - ・ AI倫理・ガバナンス・社会的課題

## ・ 試験方式

- ・ CBT(Computer Based Testing): 全国テストセンターで受験可能
- ・ 試験時間: 90分
- ・ 出題数: 約60問(多肢選択式)
- ・ 合格基準: 正答率70%前後
- ・ 実施回数: 年2回(6月・11月)

## ・ メリット

- ・ 統計・機械学習・活用事例を 広く横断的に証明できる。
- ・ 文系出身者や初学者にも取り組みやすい。
- ・ 企業(特にDX推進部門やデータ分析部門)で評価されやすい。
- ・ ITパスポートの「データサイエンス版」に近い位置づけ。

## ・ 対象者

- ・ **学生**: 就活で「データサイエンス力」をアピールしたい人
- ・ **社会人**: DX推進、データ活用部門、企画・マーケティング担当者
- ・ **技術者**: エンジニア志向だがまず基礎を固めたい人

## ・ どんな人におすすめ?

- ・ Kaggle Titanic で0.80出せる人なら → 統計検定2級+DS検定 が最も実務的に評価されやすい組み合わせ。
- ・ まだ統計や機械学習の幅広い知識を体系化していない人 → まずDS検定を受験すると知識の抜け漏れチェックになる。

# JDLA認定資格：G検定(ジェネラリスト検定)

## ・ 概要

- ・ 主催: JDLA(日本ディープラーニング協会)
- ・ 開始: 2017年
- ・ 目的: AI・ディープラーニングの基礎知識を社会に広める
- ・ 特徴: 数学やプログラミング不要、AIリテラシーの証明に最適

## ・ 試験情報

- ・ 形式: オンライン受験(自宅PC可)
- ・ 試験時間: 120分
- ・ 出題数: 220問前後(多肢選択式)
- ・ 合格率: 約60～70%

## ・ 試験範囲:

- ・ AIの歴史・基礎知識
- ・ 機械学習・ディープラーニングの代表手法(CNN, RNN, Transformerなど)
- ・ 法律・倫理・社会課題
- ・ 産業応用(医療、金融、製造など)

## ・ 対象者

- ・ 文系・理系問わず「AIを理解したい人」
- ・ 企業のDX推進担当者
- ・ エンジニア以外のビジネス職

# JDLA認定資格:E資格(エンジニア資格)

## ・ 概要

- ・ 主催: JDLA
- ・ 開始: 2018年
- ・ 目的: ディープラーニング実装スキルの証明
- ・ 特徴: 実務エンジニア・研究者向け。Python実装・数学知識必須

## ・ 試験情報

- ・ 受験条件: JDLA認定プログラム(数十時間の講座)修了必須
- ・ 形式: 会場試験(CBT)
- ・ 試験時間: 120分
- ・ 出題数: 約100問(選択式+一部計算)
- ・ 合格率: 約60%前後

## ・ 試験範囲:

- ・ 数学(線形代数、確率・統計、微分積分)
- ・ 機械学習の理論(SVM, 決定木, アンサンブルなど)
- ・ ディープラーニング(CNN, RNN, GAN, Transformer)
- ・ 実装(PyTorch/TensorFlow)

## ・ 対象者

- ・ AIエンジニア
- ・ データサイエンティスト(実装寄り)
- ・ 大学院で機械学習・AI研究を行う人

# G検定とE資格の違い(まとめ)

| 項目   | G検定           | E資格             |
|------|---------------|-----------------|
| 位置づけ | AIリテラシー       | AI実装エンジニア       |
| 受験資格 | 誰でも可          | JDLA認定講座の修了必須   |
| 試験方式 | オンライン(在宅)     | 会場CBT           |
| 出題内容 | AI全般知識・応用事例   | 数学・機械学習・深層学習の実装 |
| 難易度  | 易しめ(社会人教養レベル) | 高め(大学院・実務レベル)   |
| 主な対象 | ビジネス職・初学者     | エンジニア・研究者       |

まずAIの全体像を知りたい → G検定

Kaggle Titanicで 0.80 出せる人には「物足りない」可能性あり。

実務でAIモデルを作りたい → E資格

KaggleでTitanic 0.80を達成した人なら、認定プログラムで数理部分を補強すれば十分到達可能。

# 応用情報技術者試験: AP試験(IPA)

## 概要

- 主催: IPA(情報処理推進機構)
- 国家資格(情報処理技術者試験区分の一つ)
- 難易度: 中級(ITエンジニアの登竜門)
- 位置づけ: 情報処理技術者試験レベル3: 中級エンジニア・SE・プロジェクトリーダー候補
  - 基本情報技術者試験(FE: レベル2)の上位資格

## 試験区分

- 午前試験(選択式・多肢選択)
  - IT基礎、アルゴリズム、ネットワーク、データベース、セキュリティ、開発手法など広範囲
- 午後試験(記述式・選択問題)
  - システム開発、ネットワーク、データベース、情報セキュリティ、経営戦略、プロジェクトマネジメントなど

## 出題範囲

- テクノロジ系
  - コンピュータシステム、アルゴリズム、ソフトウェア設計、ネットワーク、データベース、セキュリティ
- マネジメント系
  - プロジェクトマネジメント、サービスマネジメント、ITIL関連
- ストラテジ系
  - 経営戦略、システム戦略、ビジネス戦略、法務(知的財産、労働関連、IT法務)

## 試験方式

- 実施: 年2回(4月・10月) … 試験時間: 午前150分、午後150分
- 合格基準: 午前・午後とも60%以上、合格率: 約20%前後(難関資格)

## メリット

- 国家資格としての信頼度が高い
- ITエンジニアとしての「広い知識」を証明できる
- 管理職やリーダー職の登用要件になることもある
- AIや統計分野も基礎レベルで出題される

## 対象者

- ITエンジニア(開発・インフラ・セキュリティなど)
- 将来プロジェクトリーダーや管理職を目指す人
- データサイエンス領域に進む前にIT基礎力を固めたい人

## まとめると:

- 基本情報技術者試験 → 初級エンジニア向け
- 応用情報技術者試験 → 中級、広い知識と応用力を証明
- 高度情報処理技術者試験 → 上級、専門分野に特化

## Kaggle LB0.80 の人は データサイエンス分野で圧倒的有利

- AP試験では AI・統計・データ分析系の問題を武器にする
- 足りないのは ネットワーク・セキュリティ・経営分野の暗記力
- 学習効率は非常に良い → 半年以内の合格が十分狙え

# Kaggleレベル別おすすめ検定ルート

| Kaggle Titanicレベル                     | 特徴                                    | おすすめ検定                           | 狙い・効果                       |
|---------------------------------------|---------------------------------------|----------------------------------|-----------------------------|
| 入門( <b>Notebook</b> を写経できる程度)LB ~0.70 | ・前処理・EDAの基礎習得中・モデルは写経で実行できる           | - 統計検定 3級/4級- データサイエンス検定(基礎理解目的) | データの基本を体系化し、数学・統計リテラシーを固める  |
| 中級(自分で特徴量を作れる)LB ~0.75                | ・欠損補完や特徴量エンジニアリングを工夫できる・クロスバリデーションを理解 | - 統計検定 2級- データサイエンス検定™           | 実務でも通じる「統計＋機械学習基礎力」を証明      |
| 上級(0.78～0.80に到達)                      | ・アンサンブルやハイパラ調整ができる・モデル解釈や評価指標を理解      | - 統計検定 準1級- G検定- 応用情報技術者試験(AP)   | データ分析＋AIリテラシー＋IT全般知識を横断的に習得 |
| エキスパート(0.80超安定)                       | ・特殊特徴量や論文知識を活用・他人のNotebookを超える工夫が可能   | - 統計検定 1級(数理・応用)<br>- E資格(実装力強化) | 数理統計の裏付け＋深層学習実装力を武器に専門家レベルへ |

- ・ 統計検定: 理論の深さを段階的に強化
- ・ データサイエンス検定: 幅広さの証明(実務・就活に強い)
- ・ G検定: AIのリテラシーをサクッと取得
- ・ E資格: 実装力を証明(AIエンジニア向け)
- ・ 応用情報: IT全般＋経営＋セキュリティの横断知識



# 今後、生成AIによるバيفコーディングが主流になるとしたら、データサイエンス系検定に意味はあるか？

バيفコーディング（AIが自動でコード生成・補助する時代）が本格化すると、従来の「人間が手でコードを書けるか」を測る試験の意義は揺らぎます。単なるスキル確認にとどまらない方向に検定内容がシフトしていくと考えられます。

# DS検定で重要度が上がるものと下がるもの

- 生成AI普及後に残る「検定の意義」
- 基盤知識の保証
  - AIがコードを書けても、仕組みが分からなければ「正しい指示」や「結果の検証」ができません。統計検定やデータサイエンス検定は、AIの出力を批判的に評価する基礎知識の証明として価値を持ち続けます。
- ITテラシーの証明
  - G検定のような「AI ITテラシー」資格はむしろ重要度が上がります。なぜならAIが生成する結果を社会・倫理・法の観点で適切に扱う力が求められるからです。
- 専門性の裏付け
  - E資格や統計検定1級のように「数理的裏付け」や「実装理論」を問う検定は、AI時代でも専門家ポジションの区別に有効です。
- 意義が相対的に下がる部分
  - 単純な暗記系・ハンスオン実装力のみを問う試験
    - 生成AIで簡単に代替可能な「文法知識」「APIの呼び方」などは資格の意義が弱まります。
    - 今後は「自動生成をどう使いこなすか」「生成物をどう検証・改善するか」に重点が移ります。

# 今後の方向性と結論

- 今後の検定の方向性

- AI活用前提の出題「AIが出力した結果の誤りを指摘せよ」「生成AIが提案した特徴量の妥当性を評価せよ」など、AIと人間の協働を前提とした試験に変わる可能性。

- 横断力の強化

- 応用情報技術者試験(IPA)のように、統計・AIに加えてネットワーク・セキュリティ・経営戦略まで含めた「広い視点」を持つことが価値になります。

- 結論

- 生成AIの時代でも、
  - 基礎理論を保証する検定(統計検定、E資格)
  - リテラシーを保証する検定(G検定、データサイエンス検定)
- は意味を持ち続けます。ただし、「コードが書ける」だけを測る試験の価値は下がり、AIを正しく使える力を示す資格の価値が上がる 方向にシフトしていくでしょう。



# 将来の「AI前提」検定のイメージ と問題例シミュレーション

## 出題の方向性

- **AI出力の評価力**: AIが生成したモデル・コード・可視化の妥当性を見抜く問題。
- **AIへの指示力(プロンプト設計力)**: 欲しい分析を適切にAIに依頼できるかを問う。
- **エラーやバイアス検出力**: AI出力の偏り・誤りを批判的に検証する力。
- **社会実装・倫理判断**: AIの適用に伴う倫理・法務・ビジネス上の影響を問う。

# 問題例シミュレーション: AI出力の評価問題

- **問題:**

- 下記は生成AIがTitanicデータセットを分析したレポートの一部です。誤りを2つ指摘し、修正方法を説明してください。

- **AI出力の要約:**

- Age列の欠損値は「0」で補完した。
- モデル評価には「学習データのAccuracy」を使用した。

- **解答例の方向:**

- 0補完はバイアスを導入 → 「中央値/回帰補完」が妥当。
- Accuracyのみは不適切 → クロスバリデーションやF1を用いるべき。

# 問題例シミュレーション: フロント設計力問題

- **問題:**

- あなたは「顧客離反率を予測するモデル」をAIに作らせた。以下の指示文を改善しなさい。

- **元のフロント:**

- 「顧客データで予測モデルを作って」

- **解答例の方向:**

- 「CSV形式の顧客属性・契約履歴を用いて、二値分類モデル(離反:1 / 非離反:0)を構築してください。欠損値補完の方法も明示し、学習・検証データを8:2に分けてROC-AUCを評価指標としてください。」

# 問題例シミュレーション：エラーやバイアス検出問題

- ・ **問題:**

- ・ 生成AIが作ったクレジット審査モデルの重要特徴量ランキングに「性別」が含まれていました。倫理・法務の観点から、どのようなリスクがあり、どう対応すべきか述べなさい。

- ・ **解答例の方向:**

- ・ リスク: 差別的判断(ジェンダーバイアス)、法令違反の可能性。
- ・ 対応: 特徴量から性別を除外、代理変数の確認、フェアネス指標(Demographic Parityなど)を追加。

# 問題例シミュレーション：社会実装・倫理問題

- ・ **問題:**

- ・ AIが自動生成した医療診断モデルを使う場合、責任の所在は誰にあるべきか？選択肢から最も妥当なものを選べ。

1. AIベンダー
2. 病院の運営者
3. 医師本人
4. 法律で定める第三者機関

- ・ **解答例の方向:**

- ・ 「3. 医師本人」+「責任を補完する形でAIベンダー・病院も説明責任を負うべき」など複合的に論じる。

# まとめ

- **従来: コード・統計式・暗記中心**
- **未来: AIをどう使うか／どう批判的に検証するか**
- **今後、問われる力:**
  - **AIリテラシー (G検定的要素)**
  - **データ検証力 (統計検定的要素)**
  - **AI実装 + 評価力 (E資格的要素)**
  - **社会実装の理解 (応用情報的要素)**