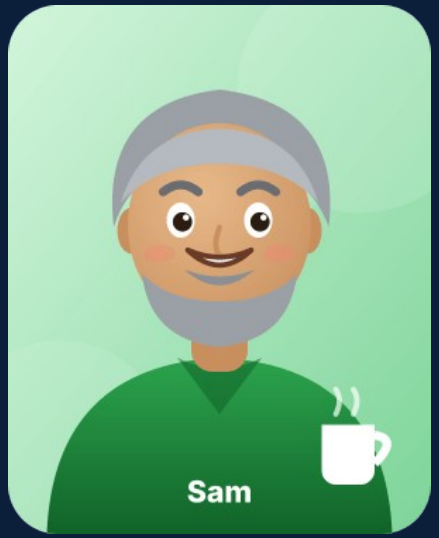
HEF-150 · MISSION 03 / 10

Hypothèses et validation

Une idée est une hypothèse qui s'ignore. La rendre testable, c'est décider à l'avance ce qui la prouverait ou la réfuterait — avant d'avoir peur du résultat.





Badge : Testeur d'idées

110 min · avancé

L'histoire

L'équipe de HashBox pense qu'un rappel par e-mail après trois jours d'inactivité augmentera la reprise des cours. Elle l'envoie à tous les membres, constate que la reprise passe de 12 % à 13,4 % et annonce un succès. Une analyste demande : « Sur combien de personnes ? Comparé à quoi ? Et si je vous montrais que d'une semaine à l'autre, la reprise varie spontanément de 11 % à 14 % ? » Personne ne sait.



« Nous avons célébré une variation qui était peut-être du bruit, dit Sam. Pour décider, il faut trois choses avant l'expérience : une hypothèse précise, un critère de succès chiffré, et un moyen de distinguer un effet réel du hasard. Sinon, on trouve toujours une raison de se féliciter. »

L'histoire

L'histoire (suite)

Tu apprends à écrire des hypothèses testables, à repérer l'hypothèse la plus risquée, à choisir une expérience adaptée (porte factice, concierge, test A/B), à calculer la taille d'échantillon, à éviter les erreurs de lecture (regarder trop tôt, tests multiples) et à décider : continuer, pivoter, arrêter.

Ta mission

À la fin, je serai capable de…

1. écrire une hypothèse falsifiable avec critère de succès fixé à l'avance
2. identifier l'hypothèse la plus risquée à tester en premier
3. choisir une expérience proportionnée (porte factice, concierge, test A/B)
4. calculer une taille d'échantillon et réaliser un test de deux proportions
5. expliquer et simuler les pièges : regarder les résultats trop tôt, comparaisons multiples

Défi d'échauffement

Défi d'échauffement

Tu penses qu'un nouveau trajet est plus rapide. Avant de le tester, quelle règle te donnes-tu pour dire « oui, il est vraiment plus rapide » : un gain de 2 minutes sur un essai ? De 5 minutes sur la moyenne de 10 essais ? Quelles raisons ordinaires (trafic, météo) pourraient fausser la comparaison ? Pourquoi fixer la règle *avant* de faire les essais ?

Le concept

Une hypothèse bien formée

```
Nous croyons que [changement]
pour [segment]
produira [résultat observable],
parce que [raison].
Nous saurons que c'est vrai quand [métrique] atteindra [seuil] en [durée].
```

Exemple : « Nous croyons qu'un rappel par e-mail après 3 jours d'inactivité, envoyé aux membres inscrits depuis moins de 30 jours, fera passer la reprise de cours à 7 jours de 12 % à 15 %, parce que l'oubli est la première cause d'abandon dans les entretiens. Nous saurons que c'est vrai quand la reprise du groupe recevant le rappel dépassera celle du groupe témoin d'au moins 2 points sur 4 semaines. »

Une hypothèse **falsifiable** peut être **réfutée** par les faits. « Les utilisateurs aimeront » n'est pas falsifiable.

Le concept

Les hypothèses sous-jacentes

Une idée repose sur plusieurs hypothèses. On les **liste**, puis on les classe par :

- **risque** : si elle est fausse, tout s'effondre ? (importance) ;
- **incertitude** : sait-on si elle est vraie ? (manque de preuves).

On teste d'abord l'hypothèse **importante et incertaine** : c'est la **plus risquée**. Les quatre catégories de risque (valeur, utilisabilité, faisabilité, viabilité — mission 1) aident à ne rien oublier.

Les expériences, du plus léger au plus lourd

Expérience	Principe	Apprend
Entretien de problème	écouter	le problème existe-t-il ?

Le concept

Les expériences, du plus léger au plus lourd (suite)

Expérience	Principe	Apprend
Porte factice (*fake door*)	afficher le bouton d'une fonction qui n'existe pas ; mesurer les clics	l'intérêt réel (comportement)
Page de destination	présenter la promesse, mesurer les inscriptions	la demande
Prototype	maquette cliquable	utilisabilité
Concierge	rendre le service **à la main** à quelques clients	la valeur, avant d'automatiser
Magicien d'Oz	l'utilisateur croit à un produit automatique, un humain le fait	idem
Test A/B	deux versions à des groupes aléatoires, comparaison d'une métrique	l'effet causal d'un changement

Règle : **le moins cher qui répond à la question**. Un test A/B de trois mois pour savoir si quelqu'un veut une fonctionnalité est un gaspillage.

Le concept

Critères de décision fixés à l'avance

Avant l'expérience, écrire :

Résultat observé	Décision
≥ seuil de succès	**persévérer**
entre seuil d'échec et de succès	**continuer à apprendre** (itérer, prolonger)
≤ seuil d'échec	**pivoter ou arrêter**

Décider **après** avoir vu le résultat conduit à s'arranger avec les chiffres.

Test A/B : les bases statistiques

Deux groupes **aléatoires** : A (témoin) et B (variante). Métrique de conversion p_A, p_B.

Le concept

Test A/B : les bases statistiques (suite)

- **Hypothèse nulle** H_0 : pas de différence réelle.
- **Test de deux proportions** (z) : $z = (p_B - p_A) / \sqrt{p(1-p)(1/n_A + 1/n_B)}$ avec p la proportion globale. $|z| > 1{,}96 \leftrightarrow$ p-valeur < 5 %.
- **Erreur de type I** (faux positif, α = 5 %) : voir un effet qui n'existe pas.
- **Erreur de type II** (faux négatif, β) : manquer un effet réel. La **puissance** $1 - \beta$ (typiquement 80 %).
- **Effet minimal détectable** : plus l'effet est petit, plus l'échantillon doit être grand.

Taille d'échantillon **par groupe** (approximation) :

```
n ≈ 2 · (z_α/2 + z_β)² · p(1−p) / δ²      (δ = différence à détecter ; z_α/2 = 1,96 ; z_β = 0,84 pour 80 % de puissance)
```

Pour détecter 12 % → 14 % (δ = 0,02) : environ **4 000 personnes par groupe**.

Le concept

Les pièges de lecture

Piège	Explication	Remède
Regarder trop tôt (*peeking*)	arrêter dès que p < 0,05 gonfle les faux positifs (jusqu'à 20-30 % avec de nombreux coups d'œil)	fixer la durée / l'échantillon à l'avance, ou utiliser des méthodes séquentielles
Comparaisons multiples	tester 20 métriques : une sera « significative » par hasard	une métrique principale décidée à l'avance, correction (Bonferroni)
Effet de nouveauté	l'intérêt initial retombe	durée suffisante
Contamination	les groupes se mélangent	randomisation par utilisateur
Échantillon trop petit	résultat bruité	calcul de taille préalable
Significativité ≠ importance	un effet minuscule peut être « significatif » avec un énorme échantillon	regarder l'**ampleur** et l'intervalle de confiance

Le concept

Quand le test A/B n'est pas possible

Peu de trafic, produit neuf : privilégier les expériences **qualitatives** et les engagements (précommandes, listes d'attente, paiement d'avance) plutôt que des statistiques sur de petits nombres.

Décider : persévérer, pivoter, arrêter

Un **pivot** change une hypothèse fondamentale (segment, problème, solution) en gardant ce qu'on a appris. **Arrêter** est une décision de succès quand elle économise des mois d'effort sur une idée réfutée.

🧠 ***Astuce** : écris la phrase « si on observe X, on arrête » avant l'expérience. Si tu n'arrives pas à l'écrire, l'expérience ne servira pas à décider.*

Exemple pas à pas

Exemple pas à pas

Problème : tester une hypothèse de rappel par e-mail avec un test de deux proportions, calculer la taille d'échantillon nécessaire, simuler la puissance, et montrer par simulation que regarder les résultats trop tôt multiplie les faux positifs.

```
import math, random

# ---------------- 1. test de deux proportions ----------------
def test_deux_proportions(succes_a, n_a, succes_b, n_b):
    pa, pb = succes_a / n_a, succes_b / n_b
    p = (succes_a + succes_b) / (n_a + n_b)
    erreur = math.sqrt(p * (1 - p) * (1 / n_a + 1 / n_b))
    z = (pb - pa) / erreur
    # p-valeur bilatérale via la fonction d'erreur
    pvaleur = math.erfc(abs(z) / math.sqrt(2))
    ecart = math.sqrt(pa * (1 - pa) / n_a + pb * (1 - pb) / n_b)
    return {"diff": pb - pa, "z": z, "p": pvaleur, "ic95": (pb - pa - 1.96 * ecart
   , pb - pa + 1.96 * ecart)}
```

Exemple pas à pas (suite)

```
# la « victoire » annoncée : 12 % → 13,4 % ; sur 1 000 personnes par groupe
r = test_deux_proportions(120, 1000, 134, 1000)
print(f"1 000 par groupe : +{r['diff'] * 100:.1f} points, z = {r['z']:.2f}, p = {r
   ['p']:.2f}, IC95 = [{r['ic95'][0] * 100:+.1f} ; {r['ic95'][1] * 100:+.1f}] poin
   ts")
assert r["p"] > 0.05 and r["ic95"][0] < 0 < r["ic95"][1]              # NON signif
   icatif : l'intervalle contient 0
# le même écart avec 20 fois plus de monde
r2 = test_deux_proportions(2400, 20000, 2680, 20000)
print(f"20 000 par groupe : +{r2['diff'] * 100:.1f} points, p = {r2['p']:.4f}, IC9
   5 = [{r2['ic95'][0] * 100:+.2f} ; {r2['ic95'][1] * 100:+.2f}] points")
assert r2["p"] < 0.001

# ---------------- 2. taille d'échantillon ----------------
def taille_echantillon(p_base, delta, alpha_z=1.96, puissance_z=0.84):
    p2 = p_base + delta
    p_moy = (p_base + p2) / 2
    return math.ceil(2 * (alpha_z + puissance_z) ** 2 * p_moy * (1 - p_moy) / delt
   a ** 2)
```

Exemple pas à pas (suite)

```
print("effet à détecter (12 % de base) | personnes par groupe")
for delta in (0.05, 0.03, 0.02, 0.01):
    print(f"        +{delta * 100:.0f} point(s)           |  {taille_echantillon(
   0.12, delta):6d}")
assert taille_echantillon(0.12, 0.02) > 3500 and taille_echantillon(0.12, 0.01) >
   3.5 * taille_echantillon(0.12, 0.02)      # moitié de l'effet → ~4× plus de mon
   de

# ---------------- 3. puissance par simulation ----------------
def puissance(p_base, delta, n, essais=1500, graine=1):
    rng, gagnes = random.Random(graine), 0
    for _ in range(essais):
        a = sum(rng.random() < p_base for _ in range(n)); b = sum(rng.random() < p
   _base + delta for _ in range(n))
        if a + b and test_deux_proportions(a, n, b, n)["p"] < 0.05: gagnes += 1
    return gagnes / essais
n = taille_echantillon(0.12, 0.02)
print(f"puissance mesurée avec n = {n} par groupe pour un effet réel de +2 points
   : {puissance(0.12, 0.02, n):.0%}  (visé : 80 %)")
```

Exemple pas à pas (suite)

```
print(f"puissance avec seulement n = 1000 : {puissance(0.12, 0.02, 1000):.0%}  → l
   'effet réel est manqué la plupart du temps")
assert 0.72 < puissance(0.12, 0.02, n) < 0.88 and puissance(0.12, 0.02, 1000) < 0.
   35

# ---------------- 4. regarder trop tôt (peeking) : des tests A/A, SANS aucune dif
   férence réelle ----------------
def faux_positifs(regarde_tout_le_temps, essais=1000, n_max=3000, pas=200, graine=
   2):
    rng, faux = random.Random(graine), 0
    for _ in range(essais):
        a = b = 0; vu = 0
        for n in range(1, n_max + 1):
            a += rng.random() < 0.12; b += rng.random() < 0.12                # l
   es DEUX groupes ont la même vraie conversion
            if regarde_tout_le_temps and n % pas == 0 and test_deux_proportions(a,
    n, b, n)["p"] < 0.05:
                vu = 1; break                                                  #
    on arrête dès que c'est « significatif »
```

Exemple pas à pas (suite)

```
        if not regarde_tout_le_temps and test_deux_proportions(a, n_max, b, n_max)
   ["p"] < 0.05: vu = 1
        faux += vu
    return faux / essais
fp_fixe, fp_peeking = faux_positifs(False), faux_positifs(True)
print(f"faux positifs : analyse unique à la fin = {fp_fixe:.1%} ; regarder toutes
   les 200 personnes = {fp_peeking:.1%}")
assert fp_fixe < 0.09 and fp_peeking > 2.5 * fp_fixe

# ---------------- 5. comparaisons multiples ----------------
def au_moins_un_faux(k, alpha=0.05): return 1 - (1 - alpha) ** k
print("métriques testées | probabilité d'au moins un faux positif | seuil de Bonfe
   rroni")
for k in (1, 5, 10, 20):
    print(f"        {k:2d}        |             {au_moins_un_faux(k):5.1%}
          |    {0.05 / k:.4f}")
assert round(au_moins_un_faux(20), 2) == 0.64

# ---------------- 6. règle de décision fixée À L'AVANCE ----------------
```

Exemple pas à pas (suite)

```
def decision(ic95, seuil_succes=0.02, seuil_echec=0.005):
    bas, haut = ic95
    if bas >= seuil_succes: return "persévérer"
    if haut <= seuil_echec: return "arrêter ou pivoter"
    return "continuer à apprendre"
print(decision(r["ic95"]), "|", decision(r2["ic95"]), "|", decision((0.025, 0.045)
   ), "|", decision((-0.004, 0.003)))
assert [decision(r["ic95"]), decision((0.025, 0.045)), decision((-0.004, 0.003))]
   == ["continuer à apprendre", "persévérer", "arrêter ou pivoter"]
```

Lecture :

1. Avec 1 000 personnes par groupe, l'écart de **+1,4 point** n'est **pas significatif** (p ≈ 0,5 ; l'intervalle contient 0) : l'annonce de succès n'était pas justifiée. Avec 20 000 par groupe, le même écart est hautement significatif.

Exemple pas à pas (suite)

2. La **taille d'échantillon** grandit comme l'inverse du carré de l'effet : détecter +1 point exige près de **4 fois plus** de monde que +2 points.
3. La **simulation de puissance** confirme le calcul : avec la taille prévue, on détecte l'effet réel environ 80 % du temps ; avec 1 000 par groupe, on le manque presque toujours.
4. Le **peeking** : sur des tests A/A (aucune vraie différence), regarder toutes les 200 personnes et s'arrêter dès que « p < 0,05 » multiplie par plus de 3 la fréquence des faux positifs.
5. Tester **20 métriques** donne 64 % de chances d'en trouver une « significative » par pur hasard.
6. La **règle de décision** est écrite avant l'expérience et s'applique à l'intervalle de confiance.

À toi, avec aide

On veut tester une page d'inscription simplifiée. Conversion actuelle : **8 %**. On ne s'intéresse à la modification que si elle apporte au moins **+1,5 point**. Le site reçoit **600 visiteurs par jour**, répartis également entre les deux versions.

1. Combien de visiteurs par groupe faut-il (puissance 80 %, α = 5 %) ? Combien de jours ?
2. Après 5 jours, on observe A : 118/1500, B : 141/1500. Calcule la p-valeur et l'intervalle de confiance. Que décides-tu ?
3. Si le trafic était de 100 visiteurs par jour, que ferais-tu à la place d'un test A/B ?

À toi, avec aide · réponse

Indices progressifs

Indice 1

Utilise `taille_echantillon(0.08, 0.015)`.

Indice 2

Utilise `test_deux_proportions`.

Indice 3

Avec peu de trafic, l'effet détectable est grand : cherche des expériences qualitatives.

Réponse

```
import math
def taille(p, d, za=1.96, zb=0.84):
    pm = p + d / 2
    return math.ceil(2 * (za + zb) ** 2 * pm * (1 - pm) / d ** 2)
def test(sa, na, sb, nb):
    pa, pb = sa / na, sb / nb; p = (sa + sb) / (na + nb)
    z = (pb - pa) / math.sqrt(p * (1 - p) * (1 / na + 1 / nb)); e = math.sqrt(pa *
    (1 - pa) / na + pb * (1 - pb) / nb)
    return z, math.erfc(abs(z) / math.sqrt(2)), (pb - pa - 1.96 * e, pb - pa + 1.9
   6 * e)
n = taille(0.08, 0.015)
print("par groupe :", n, "→ jours :", math.ceil(2 * n / 600))
z, p, ic = test(118, 1500, 141, 1500)
print(f"z = {z:.2f}, p = {p:.3f}, IC95 = [{ic[0] * 100:+.2f} ; {ic[1] * 100:+.2f}]
    points")
assert 5000 < n < 6000 and p > 0.05
```

1. **≈ 5 600 visiteurs par groupe** (≈ 11 100 au total), soit **19 jours** à 600 visiteurs par jour.

Réponse (suite)

2. À 5 jours : B = 9,4 % contre A = 7,9 % (+1,5 point), mais p ≈ 0,13, l'intervalle contient 0 : **on ne conclut pas**. Arrêter maintenant parce que l'écart « ressemble » au gain visé serait du *peeking*. On poursuit jusqu'à la taille prévue.
3. Avec 100 visiteurs par jour, le test A/B durerait 90 jours : mieux vaut un **test d'usage qualitatif** (5 personnes devant l'ancien et le nouveau formulaire, observer où elles hésitent) ou un changement radical (effet plus grand, donc détectable avec moins de monde).

À toi, en solo

Test séquentiel contrôlé. Pour pouvoir regarder les résultats pendant l'expérience sans gonfler les faux positifs :

1. implémente une règle d'arrêt à `K = 5` analyses intermédiaires avec un seuil **ajusté** (Pocock : p < 0,0158 à chaque analyse pour α global = 5 %, approximation à utiliser) ;
2. simule 2 000 tests A/A (aucune différence) : mesure le taux de faux positifs avec (a) seuil 0,05 à chaque coup d'œil, (b) seuil ajusté 0,0158, (c) une seule analyse finale à 0,05 ;
3. simule 2 000 tests A/B avec un vrai effet de +2 points (conversion 12 % → 14 %, n_max = taille calculée) : mesure la puissance de chaque règle et le nombre moyen de visiteurs utilisés avant l'arrêt ;
4. conclus : quel compromis le test séquentiel apporte-t-il ?

À toi, en solo · réponse

Pistes de correction

```
import math, random

def pval(sa, na, sb, nb):
    pa, pb = sa / na, sb / nb; p = (sa + sb) / (na + nb)
    if p in (0, 1): return 1.0
    z = (pb - pa) / math.sqrt(p * (1 - p) * (1 / na + 1 / nb)); return math.erfc(a
   bs(z) / math.sqrt(2))

def essai(p_a, p_b, n_max, regle, graine, K=5):
    rng = random.Random(graine); a = b = 0
    points = [n_max * (k + 1) // K for k in range(K)]
    for n in range(1, n_max + 1):
        a += rng.random() < p_a; b += rng.random() < p_b
        if n in points:
            if regle == "naif" and pval(a, n, b, n) < 0.05: return True, 2 * n
            if regle == "pocock" and pval(a, n, b, n) < 0.0158: return True, 2 * n
            if regle == "final" and n == n_max: return pval(a, n, b, n) < 0.05, 2
   * n
    return False, 2 * n_max
```

À toi, en solo · réponse

Pistes de correction (suite)

```
def mesurer(p_b, regle, essais=2000, n_max=3000):
    positifs = total = 0
    for g in range(essais):
        ok, n = essai(0.12, p_b, n_max, regle, g); positifs += ok; total += n
    return positifs / essais, total / essais

print("règle           | faux positifs (A/A) | puissance (+2 pts) | visiteurs moye
   ns (effet réel)")
resultats = {}
for regle in ("naif", "pocock", "final"):
    fp, _ = mesurer(0.12, regle); pu, visiteurs = mesurer(0.14, regle)
    resultats[regle] = (fp, pu, visiteurs)
    print(f"{regle:15s} |       {fp:6.1%}       |      {pu:6.1%}       |        {v
   isiteurs:7.0f}")
assert resultats["naif"][0] > 0.10 and resultats["pocock"][0] < 0.075 and resultat
   s["final"][0] < 0.075
assert resultats["pocock"][2] < resultats["final"][2]                    # le sé
   quentiel s'arrête souvent plus tôt quand l'effet est réel
```

Pistes de correction (suite)

Le seuil **naïf** gonfle les faux positifs (≈ 2 à 3 fois plus que 5 %). Le seuil de **Pocock** les maintient proches de 5 % tout en autorisant les arrêts anticipés quand l'effet est réel (moins de visiteurs en moyenne). Le prix : une puissance un peu plus faible qu'une analyse finale unique à effectif égal. On échange donc un peu de puissance contre la **possibilité de s'arrêter plus tôt**, honnêtement.

Pièges classiques

Piège	Pourquoi c'est un problème	Comment l'éviter
Hypothèse floue (« ils aimeront »)	Impossible à réfuter, donc à apprendre	Format Nous croyons / Nous saurons que c'est vrai quand
Décider du critère après avoir vu les résultats	On s'arrange avec les chiffres	Seuils de succès et d'échec fixés avant
S'arrêter dès que p < 0,05	Faux positifs multipliés par 3 ou plus	Taille et durée fixées, ou méthode séquentielle
Tester 20 métriques et célébrer celle qui sort	Un faux positif quasi certain	Une métrique principale décidée à l'avance, correction

Mini-boss

Plan d'expérimentation de HashBox. Pour l'idée « rappel par e-mail après 3 jours d'inactivité » :

1. liste au moins 6 hypothèses sous-jacentes (valeur, utilisabilité, faisabilité, viabilité) et classe-les par risque × incertitude ; désigne la plus risquée ;
2. propose une expérience **proportionnée** pour chacune des trois plus risquées (porte factice, concierge, A/B) ;
3. écris l'hypothèse principale au format complet, la métrique principale, le segment, la durée, la taille d'échantillon calculée et les règles de décision (persévérer / apprendre / arrêter) ;
4. simule l'expérience (graine fixée) avec trois « mondes » : effet réel +3 points, +0,5 point, 0 ; applique ta règle de décision et ta règle séquentielle ; mesure le taux de bonnes décisions sur 500 répétitions ;
5. calcule le coût de l'expérience en jours de trafic et le compare à celui de construire la fonctionnalité sans test (coût hypothétique) ;

Mini-boss

Mini-boss (suite)

6. rédige le compte rendu d'une expérience qui **réfute** l'hypothèse et la décision d'arrêter, sans chercher à sauver l'idée.

Mini-boss · réponse

Pistes de correction

Un bon travail : tableau des hypothèses (par exemple « les membres ouvrent les e-mails de la plateforme » est important et incertain, donc testable par porte factice ou envoi à 200 personnes), expériences adaptées, plan chiffré (taille, durée), simulation des trois mondes avec un taux de bonnes décisions élevé quand l'effet est réel et très faible taux de faux « persévérer » quand il est nul, comparaison coût de l'expérience contre coût de développement, et compte rendu honnête d'une réfutation (« nous arrêtons : l'intervalle de confiance exclut +1 point »).

Quiz éclair

Quiz éclair

1. Une hypothèse falsifiable est une hypothèse :

A. qu'on ne peut jamais tester
B. que des observations peuvent réfuter, avec un critère fixé à l'avance
C. qui plaît à tout le monde

2. S'arrêter dès que « p < 0,05 » pendant l'expérience :

A. n'a aucun effet
B. réduit les faux positifs
C. augmente fortement le taux de faux positifs

3. Pour détecter un effet deux fois plus petit, l'échantillon nécessaire est environ :

A. quatre fois plus grand

Quiz éclair

Quiz éclair (suite)

B. deux fois plus grand

C. identique

4. 🔁 *Révision mission 02* : Pour obtenir une information fiable en entretien, on demande plutôt :

A. si elle utiliserait notre idée

B. ce que la personne a fait la dernière fois

C. si elle trouve l'idée géniale

5. 🔁 *Révision mission 02* : L'opportunité d'un besoin est forte quand :

A. l'importance est faible

B. la satisfaction est maximale

C. l'importance est forte et la satisfaction actuelle faible

Quiz éclair · réponse

Réponses

1. **B**.
2. **C**.
3. **A**.
4. **B**.
5. **C**.

Je réfléchis

1. Quelle décision récente de ton équipe repose sur une hypothèse que personne n'a écrite ?
2. Quel seuil d'échec as-tu déjà fixé avant une expérience, et qu'as-tu fait en le voyant atteint ?

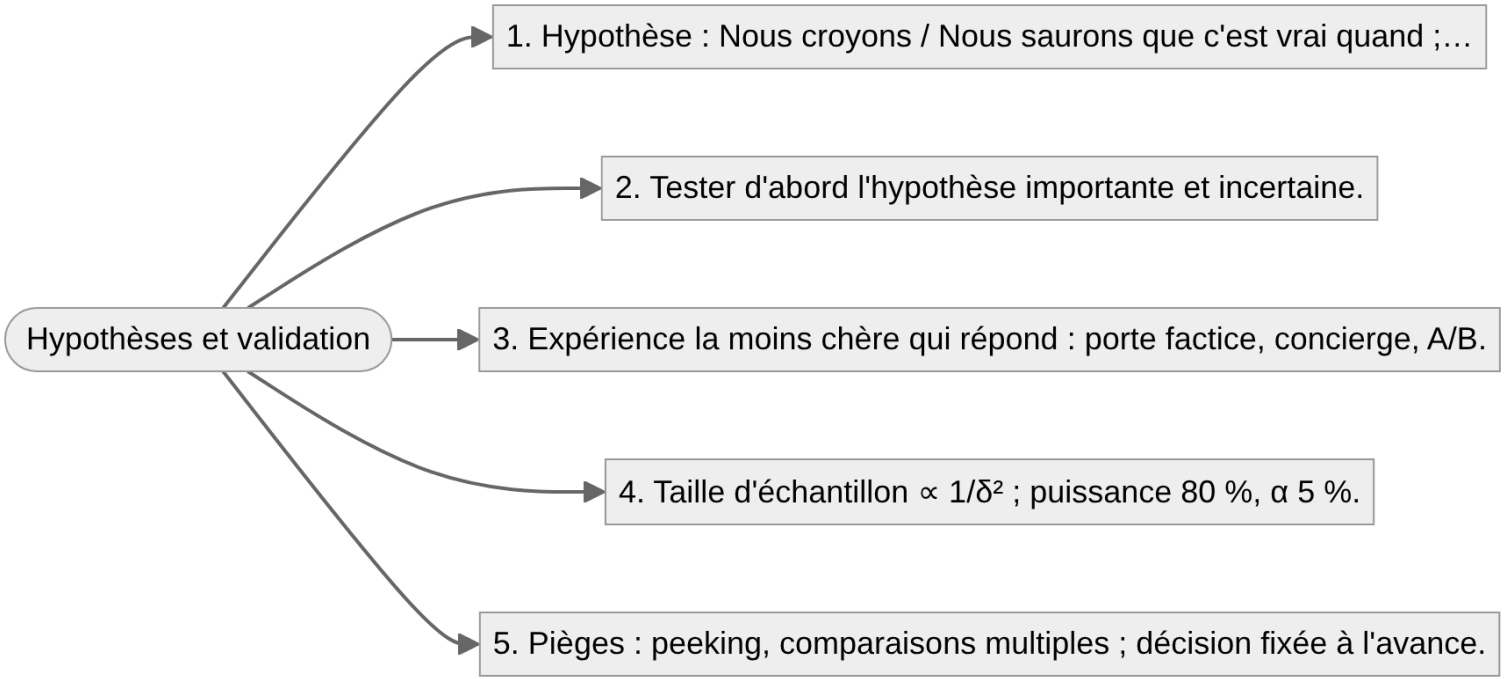
À retenir

À retenir

1. Hypothèse : Nous croyons / Nous saurons que c'est vrai quand ; falsifiable.
2. Tester d'abord l'hypothèse importante et incertaine.
3. Expérience la moins chère qui répond : porte factice, concierge, A/B.
4. Taille d'échantillon $\propto 1/\delta^2$; puissance 80 %, α 5 %.
5. Pièges : peeking, comparaisons multiples ; décision fixée à l'avance.

À retenir

Le schéma de la mission



MISSION ACCOMPLIE

+100 XP

Badge débloqué : Testeur d'idées

Prochaine mission : MVP et périmètre

